

1954-381  
MC:jdh

## **AAV2 REP PROTEIN FUSIONS**

### **CROSS-REFERENCE TO RELATED APPLICATION**

**[0001]** This application claims the benefit of U.S. provisional application Serial No. 60/432,258 filed December 11, 2002.

### **STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH**

**[0002]** This invention was made with government support in the form of grant no. HL60898-01A1 from the United States Department of Health and Human Services, National Institutes of Health, National Heart, Lung, and Blood Institute, and grant no. CA33572 from the Department of Health and Human Services, National Institutes of Health, the National Cancer Institute. The United States Government may have certain rights in the invention.

### **BACKGROUND OF THE INVENTION**

#### Technical Field

**[0003]** This invention relates to the field of molecular biology. In particular, the invention relates to methods and compositions of matter for promoting stable, site-specific integration of Rep-deleted recombinant adeno-associated virus (rAAV) vectors via delivery of a functional AAV Rep gene product to the necessary location by fusing it to an intercellular trafficking "cargo" protein such as herpes simplex virus (HSV) tegument protein, VP22 or fragment thereof.

### Description of the Background Art

**[0004]** Recombinant adeno-associated virus vectors have recently emerged as promising vehicles for gene transfer for a variety of reasons, including their lack of pathogenicity, wide host range, ability to transduce nonproliferating target cells, stable genomic integration, and comparatively low intrinsic immunogenicity. Genetic and sequence analyses of wild type AAV2 have demonstrated two primary open reading frames (ORFs). The left ORF is necessary for virus DNA replication, and contains two promoters at map positions 5 (p5) and 19 (p19). These promoters control expression from colinear, overlapping reading frames that arise from unspliced and spliced transcripts which produce Rep proteins of 78, 68, 52, and 40 kDa respectively. The right ORF, which is necessary for virion encapsulation, contains a single promoter at map position 40 (p40), and encodes three overlapping proteins (VP1, VP2, and VP3) with alternative translational initiation sites. The AAV coding regions are flanked by inverted terminal repeats (ITRs) which possess weak intrinsic promoter activity and are critical for DNA replication, encapsulation and host cell integration. See Berns, in "The Parvoviridae: The Viruses and Their Replication," Fields Virology, Fields, Knipe and Howley, Eds., 3<sup>rd</sup> edition, Lippincott-Raven, 1996, pp. 2173-2197; Chatterjee and Wong, "Adeno-associated virus vectors for transduction of genes encoding ribozymes," in Intracellular Ribozyme Applications: Principles and Protocols, Rossi and Couture (Eds.), Horizon Scientific Press, 1999; Wong and Chatterjee, "Parvovirus Vectors for Cancer Gene Therapy," in Cancer Gene Therapy, Lattine and Gershon, Eds., Academic Press, 2000.

**[0005]** One of the most interesting features of wild type AAV is its ability to integrate into a specific region in human chromosome 19 termed AAVS1. Kotin et al., *Proc. Natl.*

*Acad. Sci. USA*, 87:2211-2215, 1990; Samulski et al., *EMBO J.* 10:3941-3950, 1991. Mutational and deletion analyses have demonstrated that this property is mediated by Rep68/78, the product of the p5 promoter. Surosky et al., *J. Virol.* 71(10):7951-7959, 1997. Theoretically, the capacity to integrate site-specifically would be highly advantageous for rAAV vectors for several reasons. From a safety standpoint, nonrandom integration would lessen the likelihood of insertional mutagenesis. Kung et al., *Curr. Top. Microbiol. Immunol.* 171:1-25, 1991. In addition, cellular sequence flanking inserts are known to affect trans gene expression, resulting in varying levels of expression depending upon the location of insertion. Lacy et al., *Cell* 34(2):343-358, 1983. Targeted vector integration could minimize this variability of expression.

**[0006]** The rep gene has been removed from essentially all currently used rAAV vectors, both to provide a larger space for insertion of recombinant transgenes and to minimize the risks of recombinational events generating wild type AAV during the encapsulation process. Thus, although some studies indicate that integration is not totally random, rep-minus, wild type free rAAV stocks no longer integrate site specifically into AAVS1. Fisher-Adams et al., *Blood* 88:492-504, 1996; Rivadeneira et al., *Int. J. Oncol.* 12(4):805-810, 1998.

**[0007]** There is a need in the art for methods to improve the potential safety of rAAV vectors and to modify gene expression from rAAV vectors, in particular, methods which would allow site specific integration of rep-deleted rAAV vectors. Delivery of a functional AAV rep gene product to the necessary location would be of great value in achieving safer gene transfer with less unpredictable expression levels. Restoration of site-specific integration of rAAV

vectors could significantly impact upon the safety and utility of rAAV vectors for gene transfer and potential gene therapy.

#### **SUMMARY OF THE INVENTION**

**[0008]** Accordingly this invention provides a method for mediating site-specific integration of a rep-deleted rAAV vector in a cell, which comprises contacting the cell or expressing on the cell a fusion polypeptide which comprises an AAV2 Rep protein sequence of the left open reading frame of the rep gene that lacks a functional nuclear localization signal (NLS) and a VP22 polypeptide sequence that confers intercellular trafficking on the fusion polypeptide. The Rep protein may be fused at the carboxyl or amino terminus of the VP22 polypeptide and may be fused to it directly or indirectly, via a spacer of one or several amino acids. The AAV Rep protein preferably is truncated to remove amino acid residues 489, 490, 491 or 492 and the remaining carboxyl terminus of the translated Rep protein. The truncation most preferably is located at amino acid 490 or 491. DNA constructs and fusion proteins as described also form part of this invention. The invention also provides, in another embodiment, a method of increasing the level of integration of a rAAV vector in a cell comprising contacting the cell with a Rep fusion protein having a mutation in the AAV2 NLS.

#### **BRIEF DESCRIPTION OF THE FIGURES**

**[0009]** Figure 1 is a cartoon showing construction of VP22-AAV REP.

**[00010]** Figure 2 is a cartoon showing construction of AAV REP-VP22.

**[00011]** Figure 3 is a cartoon showing progressive carboxyl terminal deletions into the AAV2 Rep nuclear localization

signal constructed using polymerase chain reaction and fused to the amino terminal portion (3A) or carboxyl terminal portion (3B) of VP22.

**[00012]** Figure 4 shows a western blot demonstrating expression of the AAVRep<sub>490</sub>VP22 fusion protein following transfection.

**[00013]** Figure 5 is a series of immunofluorescence stains for fusion protein (A and B) and for DAPI (C and D).

**[00014]** Figure 6 is a flow chart showing the scheme for the analysis of intercellular protein trafficking using flow cytometry.

**[00015]** Figure 7 presents FACS analysis of trafficking of the AAVRep<sub>490</sub>VP22 fusion protein.

**[00016]** Figure 8 is a pair of photomicrographs of 293 cells stained with a fluorescein isotriocyanate (FITC)-conjugated antibody directed against the recombinant VP22(Gly)<sub>7</sub>AAV2Rep<sub>491</sub> protein showing VP22(Gly)<sub>7</sub>AAV2Rep<sub>491</sub> trafficking (A) and a DAPI stain to show all cells in the field (B).

**[00017]** Figure 9 is a Southern blot probed with an rAAV-specific probe (lane 1: 293 only; lane 2: Apap + VP22(Gly)<sub>7</sub>AAV2Rep<sub>491</sub>, #2; lane 3: Apap + VP22(Gly)<sub>7</sub>AAV2Rep<sub>491</sub>, #13; lane 4: Apap + VP22(Gly)<sub>7</sub>AAV2Rep<sub>491</sub>, #16; lane 5: Apap + VP22(Gly)<sub>7</sub>AAV2Rep<sub>491</sub>, #33; lane 6: 7374).

**[00018]** Figure 10 is a Southern blot probed with an AAVS1-specific probe (lane 1: 293 only; lane 2: Apap + VP22(Gly)<sub>7</sub>AAV2Rep<sub>491</sub>, #2; lane 3: Apap + VP22(Gly)<sub>7</sub>AAV2Rep<sub>491</sub>, #13; lane 4: Apap + VP22(Gly)<sub>7</sub>AAV2Rep<sub>491</sub>, #16; lane 5: Apap + VP22(Gly)<sub>7</sub>AAV2Rep<sub>491</sub>, #33; lane 6: 7374).

**[00019]** Figure 11 presents preliminary DNA sequence alignment analysis of a cell-vector junction sequence isolated following TA cloning of the junction fragment (SEQ ID NOS: 1-8).

[00020] Figure 12 shows a map of CWRHIVAPAP.

#### DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

[00021] Recently, a variety of peptides and proteins, such as the herpes simplex virus tegument protein VP22, have been shown to traffic intercellularly, both as native forms and as fusions with other proteins. See, for example, U.S. Patent No. 6,251,398. This invention takes advantage of this ability to deliver a functional AAV gene product to cells to promote site specific RAAV integration and gene delivery.

[00022] Several peptides and proteins, collectively termed "cargo" proteins, which are capable of trafficking intercellularly have been described. These include the *Drosophila antennaepedia* protein, the HIV-1 tat protein and herpes simplex virus (HSV) tegument protein, VP22. See Schwarze and Dowdy, *Trends Pharmacol. Sci.* 21(2):45-48, 2000. Any known cargo protein is contemplated for use in the invention. Peptides and proteins fused in frame to these cargo proteins also are transported intercellularly, and, most importantly, can retain function. Intercellular transport and nuclear accumulation in vitro have been described with VP22 fused to green fluorescent protein (GFP), the tumor suppressor protein p53, and the herpes simplex virus thymidine kinase suicide gene. Elliott and O'Hare, *Cell* 88(2):223-233, 1997; Phelan et al., *Nat. Biotechnol.* 16(5):440-443, 1998; Dilber et al., *Gene Ther.* 6(1):12-21, 1999. Analogous studies have been reported for HIV-1 tat fusions with several cell cycle regulatory proteins, including p27Kipl and p16INK4a. Nagahara et al., *Nat. Med.* 4(12):1449-1452, 1998; Gius et al., *Cancer Res.* 59(11):2577-2580, 1999. P-galactosidase fused to HIV-1 tat trafficked widely in an *in vivo* mouse model. Schwarze et

al., *Science* 285(5433):1569-1572, 1999). The exact mechanisms by which these proteins mediate intercellular transport have been difficult to elucidate, although transport mediated by HIV-1 tat appears to be receptor independent, and is more efficient when the tat fusion protein is denatured.

**[00023]** The ability of these cargo proteins to deliver functional genes was used in the present invention to promote site-specific rAAV integration and to increase the level of integration, to significantly enhance the potential safety of the gene delivery and to provide an improved method for expression. A variety of exemplary RepVP22 fusion constructs were constructed in which AAV rep or fragments thereof were linked in frame to the N- or C-terminus of VP22 within an expression plasmid (Invitrogen). These constructs were transfected into 293 cells, where protein expression, intercellular trafficking, and Rep function were monitored. These fusion constructs, for example AAV2Rep<sub>490</sub>VP22, VP22(Gly<sub>4</sub>)-AAV2Rep<sub>491</sub> and VP22(Gly<sub>7</sub>)-AAV2Rep<sub>491</sub>, are considered part of the present invention. See Tables I and III-IV below and Figure 3.

**[00024]** Fusion constructs according to this invention are designed to traffic intercellularly by eliminating interference by the NLS present in AAV2 Rep. The AAV2 Rep NLS extends from amino acids 485-519 of the translated Rep ORF. A mutation in or truncation of the gene which deletes all or part of the NLS such that the NLS function is lost restores trafficking ability. Thus, according to the invention, genes truncated or otherwise mutated to remove the protein's ability to signal for nuclear localization are useful to deliver any desired gene and to promote high levels of site-specific integration of the gene and improve expression qualitatively and quantitatively. Therefore, any fusion polypeptide or DNA construct encoding such

polypeptide having these properties may be used in the present invention.

**[00025]** Any polypeptide sequence that confers nuclear localization on the fusion polypeptide, as known in the art, may be used with the inventive compositions and methods. For example, VP22 polypeptides or fragments or variants thereof which retain the desired nuclear localization function are preferred. Other polypeptides suitable for use in these inventive fusion polypeptides include Drosophila antennapedia protein, HIV-1 tat protein and functional fragments or variants thereof. Functional segments of the polypeptide, whether truncated at the carboxyl or amino terminus, or both or internal deletions are included in the term fragment. The term variant includes polypeptides containing amino acid substitutions, whether conservative or not, which are at least 80% homologous and preferably 90%, 95% or 99% homologous to the native sequence and which retain the desired nuclear localization function. Persons of skill in the art are well aware of methods of constructing or purifying such molecules and of manipulating them by molecular biological techniques to construct the desired DNA and protein fusions.

**[00026]** Rep protein sequences encoded by the left open reading frame of the AAV2 rep gene that lack a functional nuclear localization signal sequence are suitable for use with the invention. Any such Rep protein sequence may be used, including sequences having a mutation in the NLS which disturbs the NLS function sufficiently to restore trafficking ability. Persons of skill in the art are aware of known methods for determining whether this trafficking ability or the NLS function is present, absent, or sufficiently reduced to allow the inventive methods to operate in the system of choice, using known or routinely modified assays and other techniques. Therefore, any AAV2



Rep protein sequence in which NLS function is absent or severely curtailed (i.e. not detectable or at a level which does not interfere with the functioning of the inventive method) compared to the activity of full-length native Rep protein is contemplated for use with this invention.

**[00027]** Specifically, Rep protein sequences in which the NLS is deleted may be used, for example by deletion of amino acids 485-519 of the native sequence or by truncation of the carboxyl terminal portion of the Rep protein at amino acid 485, amino acid 486, amino acid 487, amino acid 488, amino acid 489, amino acid 490, amino acid 491, amino acid 492, amino acid 493 or amino acid 494. By truncation at an amino acid residue, it is indicated that the amino acids carboxyl terminal to the named amino acid are removed. For example, in a protein truncated at amino acid 491, the carboxyl terminal residue of such a protein would be amino acid 491. Any deletions of the NLS which disturb function as described above may be used. For example deletion of amino acids 485-519 or 486-518 or 489-492 are suitable. Persons of skill in the art consider it routine to construct a variety of such deletion mutants and/or truncations of proteins. Therefore, such variations are considered part of the inventive compositions and methods. Rep protein mutants having point mutations in the NLS also may be used, as well as Rep protein sequences in which all or part of the NLS sequence has been removed and replaced with non-functional spacer amino acid residues.

**Table I. Description of Exemplary RepVP22 DNA Constructs.**

CONSTRUCT	DESCRIPTION
AAV2Rep <sub>490</sub> VP22	AAV2Rep truncated at amino acid 490 and fused in frame to the amino terminal end of VP22
VP22 (Gly) <sub>4</sub> AAVRep <sub>491</sub>	AAV2Rep truncated at amino acid 491 and fused in frame to the carboxyl terminal end of VP22 with DNA encoding 4 glycine residues separating the two open reading frames
VP22 (Gly) <sub>7</sub> AAVRep <sub>491</sub>	AAV2Rep truncated at amino acid 491 and fused in frame to the carboxyl terminal end of VP22 with DNA encoding 7 glycine residue separating the two open reading frames

**[00028]** Western analysis demonstrated that all RepVP22 constructs were expressed as stable protein products of expected size. Full length rep fused to VP22 did not traffic intercellularly (data not shown). A fusion gene truncated at nucleotide 490 of the AAV2 rep gene sequence did traffic intercellularly as assessed by immunofluorescence microscopy and flow cytometry. See, for example, Figure 7. In this construct, Rep<sub>490</sub>-VP22, the Rep open reading frame, truncated at amino acid residue 490 of the translated Rep protein, was fused in frame to the amino terminal end of VP22. Interestingly, an analogous VP22-Rep<sub>491</sub> fusion protein did not traffic. Insertion of 4 and 7 glycine spacers to separate the VP22 and Rep<sub>491</sub> domains and circumvent potential steric hindrance to intercellular trafficking restored the ability to traffic intercellularly. See Figure 5. One of skill in the art will readily recognize that any amino acid residue may be used as a spacer provided that the goal of reducing steric hindrance can be achieved. Therefore, spacer amino acids with small side groups are preferred.

**[00029]** A PCR assay which specifically detects vector integration into AAVS1, coupled with Southern analyses,

suggested that all three constructs described in Table I promoted site specific rAAV2 vector integration. See Figures 8-10. These types of constructs therefore form the basis of a strategy to improve both the safety and efficacy of rAAV vectors.

**[00030]** To confirm rAAV integration into the AAVS1 site by Rep<sub>490</sub>VP22, PCR products containing vector-cell junction fragments were cloned and sequenced. See Figure 11. Fusion proteins were constructed with His tags to facilitate their isolation and purification. The fusion proteins were assessed for their ability to promote site specific rAAV integration by simply applying them to cells in the form of purified Rep-VP22 fusion proteins.

**[00031]** To exploit the ability of the fusion cargo proteins to deliver functional protein domains intercellularly, the wild type and several modified AAV2 Rep gene constructs were fused in frame to the VP22 ORF both in amino- and carboxyl-terminal orientations. The fusion proteins were expressed using the highly active CMV IE promoter. Although fusions of VP22 with full length AAV2 Rep did not appear to traffic, specific Rep fusion proteins in which the NLS was truncated, for example VP22(Gly<sub>7</sub>)-AAV2Rep<sub>491</sub>, trafficked intercellularly and were capable of promoting site specific integration of recombinant RAAV vectors. See Figure 8.

**[00032]** Fusion proteins according to the invention can be expressed by plasmid DNA transfection according to any method known in the art, including calcium phosphate coprecipitation, for example. Once expressed, the fusion proteins traffic to surrounding cells via the VP22 or other intercellular trafficking protein moiety, and can mediate rAAV vector site specific integration via the AAV Rep moiety. Those of ordinary skill in the art are familiar with such methods and are able to make modifications as

desired depending on the protein fusion and cell type(s) involved. Alternatively, fusion proteins can be expressed within cells by introducing expression plasmid DNA via physical methods (lipofection, electroporation, etc.) or by using a viral vector. In addition, purified fusion protein may be applied directly to cells to promote site-specific rAAV integration. Because the constructs preferably express fusion proteins with His tags (which allow easy purification by nickel column chromatography) the proteins may be purified after production in bacteria or eukaryotic cells, and then applied directly to cells at the time of rAAV vector transduction. This increases the frequency of rAAV vector integration.

## REFERENCES

The references listed below are hereby incorporated into the specification by reference.

1. Aints et al., "Intercellular spread of GFP-VP22." *J. Gene Med.* 1(4):275-9, 1999.
2. Berns, in "The Parvoviridae: The Viruses and Their Replication," *Fields Virology*, Fields, Knipe and Howley (Eds) 3d Edition, Lippincott-Raven, 1996. pp 2173-2197.
3. Brewis et al., "Evaluation of VP22 spread in tissue culture." *J. Virol.* 74(2):1051-6, 2000.
4. Chatterjee et al., "Dual target inhibition of HIV-1 in vitro by means of an adeno-associated virus antisense vector." *Science* 258:1485-1488, 1992.
5. Chatterjee et al., "Transduction of primitive human marrow and cord blood-derived hematopoietic progenitor cells with adeno-associated virus vector." *Blood* 93:1882-94, 1999.
6. Chatterjee and Wong, "Adeno-associated Virus Vectors for Transduction of Genes Encoding Ribozymes," *Intracellular Ribozyme Applications: Principles and Protocols*, Rossi and Couture (Eds.), Horizon Scientific Press, 1999.
7. Derer et al., "Direct protein transfer to terminally differentiated muscle cells." *J. Mol. Med.* 77(8):609-13, 1999.

8. Dilber et al., "Intercellular delivery of thymidine kinase prodrug activating enzyme by the herpes simplex virus protein, VP22." *Gene Ther.* 6(1):12-21, 1999.
9. Elliott and O'Hare, "Intercellular trafficking and delivery by a herpesvirus structural protein." *Cell* 88(2):223-233, 1997.
10. Elliott and O'Hare, "Intercellular trafficking of VP22-GFP fusion proteins." *Gene Ther.* 6(1):149-251, 1999.
11. Elliott and O'Hare, "Herpes simplex virus type I tegument protein VP22 induces the stabilization and hyperacetylation of microtubules." *J. Virol.* 72(8):6448-6455, 1998.
12. Elliott and O'Hare, "Cytoplasm-to-nucleus translocation of a herpesvirus tegument protein during cell division." *J. Virol.* 74(5):2131-2141, 2000.
13. Fang et al., "Intercellular trafficking of VP22-GFP fusion proteins is not observed in cultured mammalian cells." *Gene Ther.* 5(10):1420-4, 1998.
14. Fisher-Adams et al., "Integration of adeno-associated virus vector genomes in Human CD34 cells following transduction." *Blood* 88:492-504, 1996.
15. Gius et al., "Transduced p16INK4a peptides inhibit hypophosphorylation of the retinoblastoma protein and cell cycle progression prior to activation of Cdk2 complexes in late G1." *Cancer Res.* 59(11):2577-2580, 1999.

16. Kotin and Berns, "Organization of adeno-associated virus DNA in latently infected Detroit 6 cells." *Virology* 170(2): 460-7, 1989.
17. Kotin et al., "Site-specific integration by adeno-associated virus." *Proc. Natl. Acad. Sci. USA* 87:2211-2215, 1990.
18. Kung et al., "Retroviral mutagenesis of cellular oncogenes: a review with insights into the mechanisms of insertional activation." *Curr. Top. Microbiol. Immunol.* 171:1-25, 1991.
19. Lacy et al., "A foreign beta-globin gene in transgenic mice: integration at abnormal chromosomal positions and expression in inappropriate tissues." *Cell* 34(2):343-358, 1983.
20. Nagahara et al., "Transduction of full-length TAT fusion proteins into mammalian cells: TAT-p27Kipl induces cell migration." *Nat. Med.* 4(12):1449-1452, 1998.
21. Phelan et al., "Intercellular delivery of functional p53 by the herpesvirus protein VP22." *Nat. Biotechnol.* 16(5):440-443, 1998.
22. Podsakoff et al., "Stable and efficient gene transfer into non-dividing cells by adeno-associated virus (AAV)-based vectors." *J. Virol.* 68:5656-5666, 1994.
23. Rivadeneira et al., "Sites of recombinant adeno-associated virus integration." *Int. J. Oncol* 12(4):805-810, 1998.

24. Rinaudo et al., "Conditional Site-Specific Integration into Human Chromosome 19 by Using a Ligand-Dependent Chimeric Adeno-Associated Virus/Rep Protein." *J. Virol.* 74:281-294, 2000.
25. Samulski et al., "Targeted integration of adeno-associated virus (AAV) into human chromosome 19." *EMBO J.* 10:3941-3950, 1991.
26. Schwarze et al., "In vivo protein transduction: delivery of a biologically active protein into the mouse." *Science* 285(5433):1569-1572, 1999.
27. Schwarze and Dowdy, "In vivo protein transduction: intracellular delivery of biologically active proteins, compounds and DNA." *Trends Pharmacol. Sci.* 21(2):45-48, 2000.
28. Surosky et al., "Adeno-associated virus Rep proteins target DNA sequences to a unique locus in the human genome." *J. Virol.* 71(10):7951-7959, 1997.
29. Wong and Chatterjee, "Parovirus Vectors for the Cancer Gene Therapy," *Cancer Gene Therapy*, Lattime and Gershon (Eds.), Academic Press, 2000.



## EXAMPLES

### Example 1. AAVS1 Site Specific Integration of rAAV.

**[00033]** Plasmids pVP22/myc-His and pVP22/myc-His-2 were obtained from Invitrogen (Carlsbad, CA). See Figures 1 and 2. The full length AAV2 rep gene product was amplified by PCR and inserted into the pVP22/myc-His vector as an EcoRV and XbaI fragment. The AAV2 Rep68/78 open reading frame was amplified from pTZAAV, a pUC-based phagemid containing the full length, infectious AAV2 genome inserted as a Bgl II fragment.

**[00034]** To construct a VP22-Rep fusion with the full length AAV2 Rep, see Figures 1 and 2, the rep gene was amplified as a 1.9 kb fragment using the sense primer 5' GGGAGGTTT**GATATC**GCAGCCGCCATGCCGGGG 3' (SEQ ID NO:1) with incorporation of an Xba I site (bold) and the antisense primer 5' GATTTAAT**CTAGAT**ATTGTTCAAAGATGCAG 3' (SEQ ID NO:2) with incorporation of an Xba I site (bold). The rep stop codon was modified from TAA to TAT as a part of the Xba I site to permit read-through incorporation of myc and His tags at the 3' end of the fusion protein.

**[00035]** The 5'-PCR primer used for the construction of the full length Rep-VP22 fusion was 5' GGTTTGAACGCGCAG**GATATC**ATGCCGGGG 3' (SEQ ID NO: 3) which incorporated an EcoRV site (bold). Two different full length Rep-VP22 constructs were made: (1) RepVP22cys, in which the stop codon of Rep was modified to a Cysteine residue to allow read-through of VP22 and (2) RepVP22phe, in which amino acids 620 and 621 were eliminated and residue 619 was modified from a phenylalanine to a cysteine to allow for read-through of VP22. The downstream primer for the RepVP22cys construct was 5' GCCATACCTGATTTAG**GCGGCCGC**ATTGTTCAAAGATG 3' (SEQ ID NO: 4),

while the downstream primer used to generate RepVP22phe was 5' GATTAAAATCATTTAG**GCGGCCGC**AGATGCAGTCATCCAAA 3' (SEQ ID NO: 5). Both primers incorporated a Not I site (bold) for cloning purposes. See Figure 3.

**[00036]** Progressive carboxyl terminal deletions into the AAV2 NLS were constructed using polymerase chain reaction and fused to the amino terminal portion (Figure 3A) or the carboxyl terminal portion (Figure 3B) of VP22. Initial studies indicated that the AAV2REP<sub>490</sub>-VP22 fusion protein trafficked between cells, but that the corresponding VP22-AAV2REP<sub>491</sub> did not. To circumvent possible steric interference with trafficking, 4 and 7 glycines were inserted in frame between the VP22 and AAV2REP<sub>491</sub> open reading frames. All constructs were sequenced to ensure that no mutations were inadvertently introduced following PCR amplification.

**[00037]** The full-length AAV Rep fusion protein constructs were tested for their ability to traffic intercellularly as follows. 293 or COS cells were transfected with expression plasmids encoding the fusion constructs and serially examined for spread of the fusion protein using indirect immunofluorescent microscopy after staining with an antibody directed against the myc tag common to all the fusion proteins. The constructs containing the full length AAV2 Rep did not traffic.

**[00038]** pVP22-Rep constructs with truncations in the NLS were constructed in a similar fashion to the previously described full length rep constructs. Rep proteins truncated at the carboxyl end at amino acids 484 (VP22AAVRep<sub>484</sub>), 491 (VP22AAVRep<sub>491</sub>), and 519 (VP22AAVRep<sub>519</sub>) were generated by PCR cloning. For these modified proteins, the 5' end of the rep open reading frame was amplified with

the same sense primer as VP22-Rep (5' GGGAGGTTT**GATATC**GCAGCCGCCATGCCGGGG 3'; SEQ ID NO: 1) and incorporated an EcoRV site (bold). The 3' end of the Rep ORF for 484, 491 and 519 truncations were amplified with antisense primers, 5' GGCTCCACCCTTTTGT**TCTAGAA**ATTCATGCTCCAC 3' (SEQ ID NO: 6), 5' GGGGGCGGGTCTTT**TCTAGAG**CTCCACCCTTTTGT 3' (SEQ ID NO: 7), and 5' GTTGATCGAAGCT**TCTAGAT**CTGACGTCGATGG 3' (SEQ ID NO: 8), respectively, all of which incorporated an Xba I site (bold).

**[00039]** For VP22(Gly)<sub>4</sub>AAVRep<sub>491</sub> and VP22(Gly)<sub>7</sub>AAVRep<sub>491</sub> constructs, the 5' end of Rep ORF was amplified with 5' CCATTTTGAAGC**GATATCGGTGGAGGCGGAGCCGCCATGCCGGGG** 3' (SEQ ID NO:9) and 5' GGGTCTCCATTT**GATATCGGGGGGGGTGGAGGCGGAGGC** GCCATGCCGGGG 3'(SEQ ID NO:10), respectively. EcoRV sites are in bold while bases encoding the glycine spacer residues are in bold and italicized. For the 3' end, the antisense primer for the pVP22-Rep491 protein (SEQ ID NO:7) was used. The amplified products were digested with EcoRV and XbaI, and inserted into similarly digested pV22/myc-His. Two full-length RepVP22 and three truncated RepVP22 constructs were generated.

**[00040]** Three truncated Rep constructs, AAVRep<sub>469</sub>VP22, AAVRep<sub>490</sub>VP22 and AAVRep<sub>505</sub>VP22, were created using independent Not I site-containing downstream primers coupled with the identical primer used to generate the full-length construct. The AAVRep<sub>469</sub>VP22 3' primer, 5' GATCCTTTGCCAG**CGGCCGCCAGTCTTTGACTTCCTGCTTGG** 3' (SEQ ID NO:11) extended from +1385 to +1425 with base changes at +1405 to +1408 and +1412. These sequence changes modified residue 469 from a phenylalanine to a cysteine and eliminated the production of all amino acids C-terminal to residue 469.

**[00041]** AAVRep<sub>490</sub>VP22 C-terminal primer, 5'

GGTCTTTT**GCGGCCGCC**ACCCTTTTG 3' (SEQ ID NO:12), extended from +1457 to +1483. Mismatches at +1469, +1471, and +1473 to +1475 were used to eliminate all residues C-terminal to 490.

AAVRep<sub>505</sub>VP22 3' primer, 5'

GACTCGCGCAC**GCGGCCG**CGCTCACTTATATCTGCG 3' (SEQ ID NO:13), extended from + 1496 to +1531. It contained nucleotide changes at positions +1513, +1515 to +1517 and +1520 resulting in the loss of amino acids C-terminal to residue 505. Additionally, residue 505 was modified from a proline to an arginine. All C-terminal primers above are given in the reverse orientation. Not I sites are indicated in bold.

**[00042]** The Rep protein sequence of the Rep<sub>491</sub> truncated construct ends at amino acid 491 of the translated Rep protein, however there are 8 amino acids at the junction leading to the initiation codon of the VP22 polypeptide sequence. These amino acids (DIQHSGGR; SEQ ID NO:14) result from additional nucleotides found within the multiple cloning site. Therefore, it is clear to one of ordinary skill that multiple variations of the fusion peptides are possible, depending on the exact construction methods used to create them. The two moieties of the fusion polypeptide may be fused directly or indirectly, with additional amino acids present at the junction or either terminus. See Table II, below for exemplary sequences contained in the Rep fusion polypeptides compared to Rep wild type. All constructs were analyzed by DNA sequencing to insure that no additional mutations were inadvertently incorporated during the PCR amplifications. See Tables III-VI for sequence information for exemplary constructs.

Table II. Sequence Comparison: Wild Type Rep and Truncated Rep-VP22.		
Name	Sequence	SEQ ID NO:
A. Rep 78 WT Rep 78-Cys-VP22	CDLVNVDLDDCIFEQ (607-621)	15
	CDLVNVDLDDCIFEQCGR-VP22	16
B. Rep 78 WT Rep <sub>491</sub> VP22	YVKKGGAKKRPAPSD (485-499)	17
	YVKKGGAADIQHSGGR-VP22	18
C. Rep 78 WT Rep <sub>505</sub> VP22	PAPSDADISEPKR (495-507)	19
	PAPSCADISERGR-VP22	20

**[00043]** All rep gene inserts were amplified using a PE 9600 thermal cycler (Perkin and Elmer). A standard 100 µl reaction contained 100 ng of template DNA, 25 pmol of each respective upstream and downstream primer, 2 units of Vent polymerase (New England Biolabs, Beverly, MA), 200 µM of each dNTP, 3mM MgSO<sub>4</sub>, and 1x Vent reaction buffer. The mixture was denatured at 95°C for 5 minutes, and then 25 cycles of amplification (95°C, 30 s; 60°C, 30 s; 72°C, 90 s) were performed, followed by one extension cycle at 72°C for 7 minutes. PCR products were gel purified using the Prep-a-Gene™ purification kit (Bio-Rad Laboratories, Hercules, CA), digested with appropriate restriction enzymes (NEB) and ligated into corresponding vectors at 16°C for 16 hours. Plasmid constructs were transformed into chemically competent DH5a cells using standard methods. Plasmids were purified by anion exchange column chromatography (Qiagen, Valencia, CA), and quantitated spectrophotometrically. Enzymes were used according to conditions suggested by the manufacturers. Oligonucleotides were synthesized using a 394 B DNA Synthesizer (Applied Biosystems, Foster City, CA). All constructs were sequenced to insure that mutations were not inadvertently introduced during amplification.

**[00044]** Figure 12 shows a map of CWRHIVAPAP. This construct contains one expression cassette encoding an antisense RNA complementary to the HIV TAR region under RSV LTR control, and another cassette encoding hu-placental alkaline phosphatase (hu PLAP) under PGK promoter control. CWRPGKH is similar to CWRHIVAPAP except for substitution of a PGK hygromycin resistance cassette for the PGK PLAP cassette.

**[00045]** African green monkey Vero (#CCL-81), 293 cells, COS cells and a Detroit 6-derived cell line, 7374, which contains integrated wild type AAV2, were maintained in high glucose Dulbecco's MEM (DMEM) with 2 mM glutamine and 10% heat inactivated fetal calf serum, at 37°C in 5% humidified CO<sub>2</sub>. All cells were routinely tested and found free of mycoplasma. All transfections were performed using a CellPfect Transfection kit (calcium phosphate procedure; Amersham Pharmacia, Piscataway, NJ) according to the manufacture's directions. For Western blot of VP-Rep fusion proteins, 293 cells were transfected with VP22-Rep or Rep-VP22 constructs (or their associated modified constructs lacking a functional NLS) using calcium phosphate coprecipitation. Cells were harvested after 48 hours and lysed. Proteins were separated using SDS-PAGE electrophoresis, and transferred to nitrocellulose. The western analyses demonstrated expression of the AAVREP<sub>490</sub>VP22 fusion protein following transfection. See Figure 4.

Example 2. Rep Expression and Trafficking  
Analyses by Immunofluorescence Microscopy.

**[00046]** Amino and carboxyl-terminal VP22/AAV2 Rep fusion proteins encoded by expression plasmids were initially tested for their ability to traffic intercellularly after calcium-phosphate transfection into 293 cells. For

immunofluorescence assays, approximately  $6.0 \times 10^5$  293 cells were plated on coverslips in 6-well plates and transfected with 1-3  $\mu$ g of expression vector DNA for the various Rep derivatives. At specified times post-transfection, cells were briefly washed 3 times in room temperature phosphate-buffered saline (PBS), fixed in methanol at  $-20^\circ\text{C}$  for 5 minutes, and permeabilized by incubating them in acetone for 2-5 minutes at  $-20^\circ\text{C}$ . The fixed cells were subsequently blocked with 1% BSA/1x PBS for 5 minutes at room temperature and stained with 1  $\mu$ M primary mouse monoclonal anti-rep (such as CAT# MAB6030, Maine Biotechnology, Portland, ME) or anti-c-myc antibody (such as CAT# R950-25, Invitrogen) diluted in 1% BSA/1x PBS, for 1 hour. The cells were then washed 3 times in PBS, 5 minutes each time, and incubated for 1 hour with a FITC-conjugated goat anti-mouse IgG secondary antibody (such as CAT# sc-2010; Santa Cruz Biotechnology, Santa Cruz, CA) and DAPI (4',6-diamidino-2-phenylindole; Sigma, St. Louis, MO). Following the washes, the fixed cells were briefly rinsed in sterile  $\text{dH}_2\text{O}$ , air-dried, and mounted onto glass slides using a 50% glycerol in  $\text{dH}_2\text{O}$ . All staining procedures were conducted at room temperature. Cells were photographed by epifluorescence on a Nikon Labophot-2 photomicroscope with fluorescein and DAPI filters using a Nikon Fluor 40x objective. No visible staining of the full length construct RepVP22 was seen outside of the nucleus. Therefore, it appears that Rep-NLS overrides VP22's inherent nature to traffic outside of cell.

**[00047]** The ability of VP22AAV2Rep<sub>491</sub> and VP22(Gly)<sub>4</sub>AAV2Rep<sub>491</sub> constructs to traffic intercellularly also were compared following transfection. Cells were stained for fusion protein with fluorescein isothiocyanate (FITC) and with 4'-6-diamidino-2-phenylindole-2HCl (DAPI) to

visualize the cells. See Figure 5. Panels 5A and 5B show immunofluorescent staining indicating the presence of the fusion protein. The results indicate that the 4-glycine insert construct traffics intercellularly.

Example 4. Flow Cytometry.

**[00048]** To further confirm intercellular transport of Rep-VP22 fusions, 293 cells were transfected with expression plasmids encoding the Rep-VP22 fusion proteins and analyzed by flow cytometry. See Figure 6. A separate culture of 293 cells was stained with PKH26, a vital membrane dye which permanently stains cells and is used for measuring cell division by flow cytometry. After expression of the VP22 protein for about 48 hours, these two populations of cells were mixed, incubated, and analyzed by flow cytometry using FITC conjugated anti fusion protein antibodies. Trafficking is indicated by demonstration of a cell population that stains with both PKH26 and antibody specific for VP22 in the VP22 fusion protein. The results are shown in Figure 7. Panels A, B and C indicate that discrimination of PKH26 and antibody staining for VP22 or the VP22 fusions was comparatively specific. Analysis of cells after mixing (panels D and E) shows a comparatively large population of cells that is PKH26 and either VP22 or AAVRep<sub>490</sub>VP22 double positive, indicating trafficking.

Example 5. Site-Specific Integration.

**[00049]** An rAAV vector-containing plasmid pCWRHIVAPAP and one of the relevant Rep derivatives were cotransfected using calcium phosphate into  $1.8 \times 10^6$  293 cells seeded in 60 mm dishes. 293 cells were harvested between 60 and 90 hours post-transfection and washed twice with PBS at 4°C. Cell



pellets were suspended in 100 mM NaCl, 25 mM EDTA, and 10 mM Tris, pH 8.0, with 1 µg/ml RNase A and incubated for 2 hours at 37°C. Sodium dodecyl sulfate (SDS) and Proteinase K then were added to a final concentration of 0.5% and 0.1 mg/ml, respectively, and the mixture was incubated overnight at 56°C. Genomic DNA was purified from the digested cell pellet material by phenol/chloroform extraction, followed by ammonium acetate/ethanol precipitation. Isolated DNA was quantified via spectrophotometric analysis. Similar experiments were performed using CWRHIVAPgkH, an rAAV vector encoding resistance to hygromycin. In these experiments, cells were grown in media supplemented with 250 µg/mL hygromycin and 400 µg/mL G418 to select for cells expressing the rAAV vector and fusion protein, respectively. Colonies resistant to both hygromycin and G418 were isolated and expanded. Genomic DNA was extracted from the cell lines as described above.

**[00050]** PCR analyses employing one primer within the vector and the other primer within AAVS1 were used to assess site-specific integration. Each 50 µl reaction contained 50 ng DNA template and 25 pmol each of a specific primer for rAAV vector and for AAVS1. Reaction mixtures were denatured at 95°C for 5 minutes, cooled to 80°C for 2 minutes (at which point the Taq polymerase was added), and then subjected to 35 cycles of amplification (94°C, 1 min; 55°C, 1 min; 72°C, 3 min), followed by a single extension cycle at 72°C for 5 minutes.

**[00051]** To confirm rAAV vector site-specific integration into AAVS1, PCR products corresponding to vector cellular junction sequences were inserted into pGEM-T vectors (Promega, Madison, WI), amplified in DH5α cells, and subjected to agarose gel sequence analysis in two

independent Southern analyses, one probed with an rAAV-specific (Figure 9) and the other with an AAVS1-specific (Figure 10) probe. PCR reactions were performed using the Taq DNA polymerase kit (Qiagen), designed to amplify DNA containing secondary structures following the manufacture's directions. Amplified products were separated using 0.8% agarose gel electrophoresis in duplicate, and transferred overnight to a nitrocellulose membrane according to methods known in the art. After cross-linking the DNA samples to the filter blot, the membrane was cut in half, each half containing a complete set of the samples to be analyzed. One blot half was hybridized with a random primed <sup>32</sup>P-labeled AAV-specific probe while the other half was hybridized with a AAVS1-specific probe. Bands that were positive with both probes indicate site-specific integration. Western blots were used to confirm the different sizes of mutants. Phosphorimaging analysis (Molecular Dynamics) was used to evaluate the extent of rAAV integration. See Figure 8.

Example 6. rAAV-Cell Junction Sequence Analysis.

**[00052]** Preliminary DNA sequence alignment analyses of cell-vector junction sequences isolated following TA cloning of the junction fragment demonstrated both vector and AAVS1 sequences, indicating site-specific integration of the vector. See Figure 11.

**Table III. PVP22 (Gly<sub>4</sub>)Rep<sub>491</sub> Sequence**

gacggatcgg	gagatctccc	gatcccctat	ggtcgactct	cagtacaatc	tgctctgatg	60
ccgcatagtt	aagccagtat	ctgctccctg	cttgtgtgtt	ggaggtcgct	gagtagtgcg	120
cgagcaaaat	ttaagctaca	acaaggcaag	gcttgaccga	caattgcatg	aagaatctgc	180
ttagggttag	gcgttttgcg	ctgcttcgcg	atgtacgggc	cagatatacg	cgttgacatt	240
gattattgac	tagttattaa	tagtaatcaa	ttacggggtc	attagttcat	agcccatata	300
tggagttccg	cgttacataa	cttacggtaa	atggcccggc	tggctgaccg	cccaacgacc	360
cccgccatt	gacgtcaata	atgacgtatg	ttcccatagt	aacgccaata	gggactttcc	420
attgacgtca	atgggtggac	tatttacggg	aaactgcca	cttggcagta	catcaagtgt	480
atcatatgcc	aagtacggcc	cctattgacg	tcaatgacgg	taaatggccc	gcctggcatt	540
atgccagta	catgacctta	tgggactttc	ctacttggca	gtacatctac	gtattagtca	600
tcgctattac	catggtgatg	cggttttggc	agtacatcaa	tgggcgtgga	tagcggtttg	660
actcacgggg	atttccaagt	ctccacccca	ttgacgtcaa	tgggagtttg	ttttggcacc	720
aaaatcaacg	ggactttcca	aaatgtcgta	acaactccgc	ccattgacg	caaatgggcg	780
gtaggcgtgt	acggtgggag	gtctatataa	gcagagctct	ctggctaact	agagaacca	840
ctgcttactg	gcttatcgaa	attaatacga	ctcactatag	ggagaccaa	gctggctagt	900
taagcttatt	atgacctctc	gccgctccgt	gaagtcgggt	ccgcgggagg	ttccgcgcga	960
tgagtacgag	gatctgtact	acaccccgtc	ttcaggtatg	gcgagtcgcc	atagtccgcc	1020
tgacacctcc	cgccgtggcg	ccctacagac	acgctcgcg	cagaggggcg	aggtccgttt	1080
cgtccagtac	gacgagtcgg	attatgccct	ctacgggggc	tcgtcttcg	aagacgacga	1140
acacccggag	gtcccccgga	cgcggcgtcc	cgtttcggg	gcggttttgt	ccgccccggg	1200
gcctgcgcgg	gcgccctccg	cacccgctgg	gtccggaggg	gccggacgca	caccaccac	1260
cgccccccgg	gccccccgaa	cccagcgggt	ggcgtctaag	gcccccgcg	ccccggcggc	1320
ggagaccacc	cgcggcagga	aatcggcca	gccagaatcc	gccgactcc	cagacgcccc	1380
cgcgctcgac	gcgccaaccc	gatccaagac	accgcgcgag	gggctggcca	gaaagctgca	1440
ctttagcacc	gcccccccaa	accccgacgc	gccatggacc	ccccgggtgg	ccggctttaa	1500
caagcgcgtc	ttctgcgccg	cggtcgggcg	cctggcgggc	atgcatgcc	ggatggcggc	1560
tgtccagctc	tgggacatgt	cgcgctccg	cacagacgaa	gacctcaacg	aactccttgg	1620
catcaccacc	atccgcgtga	cggtctgcga	gggcaaaaac	ctgcttcagc	gcgccaacga	1680
gttggtgaat	ccagacgtgg	tgcaggacgt	cgacgcggcc	acggcgactc	gaggcggttc	1740
tgcggcgtcg	cgccccaccg	agcgacctcg	agccccagcc	cgctccgctt	ctcgcgccag	1800
acggcccgtc	gagggtagcg	agctcggatc	cactagtcca	gtgtggtgga	attctgcaga	1860
tatcggtgga	ggcggagccg	ccatgccggg	gttttacgag	attgtgatta	aggtccccag	1920
cgaccttgac	gggcatctgc	ccggcatttc	tgacagcttt	gtgaactggg	tggccgagaa	1980
ggaatgggag	ttgccgccag	attctgacat	ggatctgaat	ctgattgagc	aggcaccctt	2040
gaccgtggcc	gagaagctgc	agcgcgactt	tctgacggaa	tggcgccgtg	tgagtaaggc	2100
cccgagggcc	cttttctttg	tgcaatttga	gaaggagag	agctacttcc	acatgcacgt	2160
gtcgtggaa	accaccgggg	tgaatccat	ggttttggga	cgtttcctga	gtcagattcg	2220
cgaaaaactg	attcagagaa	tttaccgcgg	gatcgagccg	actttgccaa	actggttcgc	2280
ggtcaciaag	accagaaatg	gcgccggagg	cggaacaag	gtggtggatg	agtgtacat	2340
cccccaattac	ttgctcccca	aaacccagcc	tgagctccag	tgggcgtgga	ctaatatgga	2400
acagtattta	agcgctgtt	tgaatctcac	ggagcgtaaa	cggttgggtg	cgcagcatct	2460
gacgcacgtg	tcgcagacgc	aggagcagaa	caaagagaat	cagaatccca	attctgatgc	2520
gccggtgatc	agatcaaaaa	cttcagccag	gtacatggag	ctggtcgggt	ggctcgtgga	2580
caaggggatt	acctcggaga	agcagtggtg	ccaggaggac	caggcctcat	acatctcctt	2640
caatgcggcc	tccaactcgc	ggtcccaaat	caaggctgcc	ttggacaatg	cgggaaagat	2700

tatgagcctg	actaaaaaccg	cccccgacta	cctgggtgggc	cagcagcccg	tggaggacat	2760
ttccagcaat	cggatttata	aaatTTTtga	actaaacggg	tacgatcccc	aatatgcggc	2820
ttccgtcttt	ctgggatggg	ccacgaaaaa	gttcggcaag	aggaacacca	tctggctgtt	2880
tgggcctgca	actaccggga	agaccaacat	cgcggaggcc	atagcccaca	ctgtgccctt	2940
ctacgggtgc	gtaaactgga	ccaatgagaa	ctttcccttc	aacgactgtg	tcgacaagat	3000
ggtgatctgg	tgggaggagg	ggaagatgac	cgccaaggtc	gtggagtcgg	ccaaagccat	3060
tctcggagga	agcaagggtc	gcgtggacca	gaaatgcaag	tcctcggccc	agatagaccc	3120
gactcccgtg	atcgtcacct	ccaacaccaa	catgtgcgcc	gtgattgacg	ggaactcaac	3180
gaccttcgaa	caccagcagc	cgttgcaaga	cggatgttc	aaatttgaac	tcacccgcgc	3240
tctggatcat	gactttggga	aggtcaccaa	gcaggaagtc	aaagactttt	tccggtgggc	3300
aaaggatcac	gtggttgagg	tggagcatga	attctacgtc	aaaaaggggtg	gagctctaga	3360
gggcccgcgg	ttcgaaacaa	aactcatctc	agaagaggat	ctgaatatgc	ataccggtca	3420
tcatcaccat	caccattgag	tttaaaccgc	ctgatcagcc	tcgactgtgc	cttctagtgt	3480
ccagccatct	gttgtttgcc	cctccccctg	gccttccttg	accctggaag	gtgccactcc	3540
cactgtcctt	tcctaataaa	atgaggaaat	tgcatacgcat	tgtctgagta	ggtgtcattc	3600
tattctgggg	ggtgggtggg	ggcaggacag	caagggggag	gattgggaag	acaatagcag	3660
gcatgctggg	gatgcggtgg	gctctatggc	ttctgaggcg	gaaagaacca	gctggggctc	3720
tagggggtat	ccccacgcgc	cctgtagcgg	cgcattaaagc	gcggcgggtg	tggtgggttac	3780
gcgcagcgtg	accgctacac	ttgccagcgc	cctagcgcgc	gctcctttcg	ctttcttccc	3840
ttcctttctc	gccacgttcg	cggcctttcc	ccgtcaagct	ctaaatcggg	gcatcccttt	3900
agggttccga	tttagtgctt	tacggcacct	cgaccccaaa	aaacttgatt	agggtgatgg	3960
ttcacgtagt	gggccatcgc	cctgatagac	ggtttttcgc	cctttgacgt	tggagtccac	4020
gttctttaat	agtggactct	tgttcctaaac	tggacaacaa	ctcaacccta	tctcgggtcta	4080
ttcttttgat	ttataagggg	ttttggggat	ttcggcctat	tggttaaaaa	atgagctgat	4140
ttaacaaaaa	tttaacgcga	attaattctg	tggaatgtgt	gtcagttagg	gtgtggaaag	4200
tccccaggct	ccccaggcag	gcagaagtat	gcaaagcatg	catctcaatt	agtcagcaac	4260
cagggtgtga	aagtccccag	gctccccagc	aggcagaagt	atgcaaagca	tgcatactaa	4320
ttagtcagca	accatagtcc	cgccccctaac	tccgcccatac	ccgcccctaa	ctccgcccag	4380
ttccgcccac	tctccgcccc	atggctgact	aatTTTTttt	atttatgcag	aggccgaggc	4440
cgctcttgcc	tctgagctat	tccagaagta	gtgaggaggc	ttttttggag	gcctaggctt	4500
ttgcaaaaag	ctcccgggag	cttgatatatc	cattttcggg	tctgatcaag	agacaggatg	4560
aggatcgttt	cgcatagttg	aacaagatgg	attgcacgca	ggttctccgg	ccgcttgggt	4620
ggagaggcta	ttcggctatg	actgggcaca	acagacaatc	ggctgctctg	atgccgcgt	4680
gttccggctg	tcagcgcagg	ggcgcccggg	tctttttgtc	aagaccgacc	tgtccgggtgc	4740
cctgaatgaa	ctgcaggacg	aggcagcgcg	gctatcgtgg	ctggccacga	cgggcgttcc	4800
ttgcgcagct	gtgctcgacg	ttgtcactga	agcgggaagg	gactggctgc	tattgggcga	4860
agtgcggggg	caggatctcc	tgtcatctca	ccttgctcct	gccgagaaaag	tatccatcat	4920
ggtgatgca	atgcggcggc	tgcatacgct	tgatccggct	acctgcccac	tcgaccacca	4980
agcgaaacat	cgcatacgagc	gagcacgtac	tccgatggaa	gccggtcttg	tcgatcagga	5040
tgatctggac	gaagagcatc	aggggctcgc	gccagccgaa	ctgttcgcca	ggctcaaggc	5100
gcgcataccc	gacggcgagg	atctcgtcgt	gacccatggc	gatgcctgct	tgccgaatat	5160
catggtggaa	aatggccgct	tttctggatt	catcgactgt	ggccggctgg	gtgtggcgga	5220
ccgctatcag	gacatagcgt	tggctacccg	tgatattgct	gaagagcttg	gcggcgaatg	5280
ggctgaccgc	ttcctcgtgc	tttacggtat	cgcgcgtccc	gattcgcagc	gcatcgctt	5340
ctatcgcctt	cttgacgagt	tcttctgagc	gggactctgg	ggttcgcgaa	atgaccgacc	5400
aagcgacgcc	caacctgcca	tcacgagatt	tcgattccac	cgccgccttc	tatgaaagggt	5460
tgggcttcgg	aatcgttttc	cgggacgccg	gctggatgat	cctccagcgc	gggatctca	5520
tgtcggagtt	cttcgcccac	cccaacttgt	ttattgcagc	ttataatggt	tacaaataaa	5580

gcaatagcat	cacaaatttc	acaaataaag	catttttttc	actgcattct	agttgtggtt	5640
tgtccaaact	catcaatgta	tcttatcatg	tctgtatacc	gtcgacctct	agctagagct	5700
tggcgtaatc	atgggtcatag	ctgtttcctg	tgtgaaattg	ttatccgctc	acaattccac	5760
acaacatacg	agccggaagc	ataaagtgtg	aagcctgggg	tgccaatga	gtgagctaac	5820
tcacattaat	tgcgttgccg	tactgccccg	ctttccagtc	gggaaacctg	tcgtgccagc	5880
tgcattaatg	aatcggccaa	cgcgcgggga	gaggcggttt	gcgtattggg	cgctcttccg	5940
cttcctcgct	cactgactcg	ctgcgctcgg	tcgttcggct	gcggcgagcg	gtatcagctc	6000
actcaaaggc	ggtaatacgg	ttatccacag	aatcagggga	taacgcagga	aagaacatgt	6060
gagcaaaagg	ccagcaaaag	gccaggaacc	gtaaaaaggc	cgcgttgctg	gcgtttttcc	6120
ataggctccg	ccccctgac	gagcatcaca	aaaatcgacg	ctcaagtcag	aggtggcgaa	6180
acccgacagg	actataaaga	taccaggcgt	ttccccctgg	aagctccctc	gtgcgctctc	6240
ctgttccgac	cctgccgctt	accggatacc	tgtccgcctt	tctcccttcg	ggaagcggtg	6300
cgctttctca	atgctcacgc	tgtaggtatc	tcagttcggg	gtaggtcggt	cgctccaagc	6360
tgggctgtgt	gcacgaaccc	cccgttcagc	ccgaccgctg	cgcttatcc	ggtaactatc	6420
gtcttgagtc	caaccgggta	agacacgact	tatcgccact	ggcagcagcc	actggtaaca	6480
ggattagcag	agcgagggtat	gtaggcggtg	ctacagagtt	cttgaagtgg	tggcctaact	6540
acggctacac	tagaaggaca	gtatttggtg	tctgcgctct	gctgaagcca	gttaccttcg	6600
gaaaaagagt	tggtagctct	tgatccggca	aacaaaccac	cgctggtagc	ggtggttttt	6660
ttgtttgcaa	gcagcagatt	acgcgcagaa	aaaaaggatc	tcaagaagat	cctttgatct	6720
tttctacggg	gtctgacgct	cagtggaacg	aaaactcacg	ttaagggatt	ttggtcatga	6780
gattatcaaa	aaggatcttc	acctagatcc	ttttaaatta	aaaatgaagt	tttaaatcaa	6840
tctaaagtat	atatgagtaa	acttggtctg	acagttacca	atgcttaatc	agtgaggcac	6900
ctatctcagc	gatctgtcta	tttcgttcat	ccatagttgc	ctgactcccc	gtcgtgtaga	6960
taactacgat	acgggagggc	ttaccatctg	gccccagtg	tgcaatgata	ccgcgagacc	7020
cacgctcacc	ggctccagat	ttatcagcaa	taaaccagcc	agccggaagg	gccgagcgca	7080
gaagtgggtc	tgcaacttta	tccgcctcca	tccagtcctat	taattggtgc	cgggaagcta	7140
gagtaagtag	ttcgccagtt	aatagtttgc	gcaacgttgt	tgccattgct	acaggcatcg	7200
tgggtgcacg	ctcgctggtt	ggtatggctt	cattcagctc	cggttcccaa	cgatcaaggc	7260
gagttacatg	atcccccatg	ttgtgcaaaa	aagcggttag	ctccttcggt	cctccgatcg	7320
ttgtcagaag	taagttggcc	gcagtgttat	cactcatggt	tatggcagca	ctgcataatt	7380
ctcttactgt	catgccatcc	gtaagatgct	tttctgtgac	tgggtgagta	tcaaccaagt	7440
cattctgaga	atagtgtatg	cggcgaccga	gttgctcttg	cccggcgcta	atacgggata	7500
ataccgcgcc	acatagcaga	actttaaaag	tgctcatcat	tggaaaacgt	tcttcggggc	7560
gaaaactctc	aaggatctta	ccgctgttga	gatccagttc	gatgtaacct	actcgtgcac	7620
ccaactgata	ttcagcatct	tttactttca	ccagcgtttc	tgggtgagca	aaaacaggaa	7680
ggcaaaatgc	cgcaaaaaag	ggaataaggg	cgacacggaa	atggtgaata	ctcatactct	7740
tcctttttca	atattattga	agcatttatc	agggttattg	tctcatgagc	ggatacatat	7800
ttgaatgtat	ttagaaaaat	aaacaaatag	gggttccgcg	cacatttccc	cgaaaagtgc	7860
cacctgacgt	c					7871

(SEQ ID NO:21)

**Table IV. Rep491VP22 Sequence**

gacggatcgg	gagatctccc	gateccctat	ggtcgactct	cagtacaatc	tgctctgatg	60
ccgcatagtt	aagccagtat	ctgctccctg	cttgtgtgtt	ggaggctcgt	gagtagtgcg	120
cgagcaaaat	ttaagctaca	acaaggcaag	gcttgaccga	caattgcatg	aagaatctgc	180
ttagggttag	gcgttttgcg	ctgcttcgcg	atgtacgggc	cagatatacg	cgttgacatt	240

gattattgac	tagttattaa	tagtaatcaa	ttacggggtc	attagttcat	agcccatata	300
tggagttccg	cgttacataa	cttacggtaa	atggcccgcc	tggctgaccg	cccaacgacc	360
cccgccatt	gacgtcaata	atgacgtatg	ttcccatagt	aacgccata	gggactttcc	420
attgacgtca	atgggtggac	tatttacggt	aaactgccc	cttggcagta	catcaagtgt	480
atcatatgcc	aagtacgccc	cctattgacg	tcaatgacgg	taaatggccc	gcctggcatt	540
atgccagta	catgacctta	tgggactttc	ctacttggca	gtacatctac	gtattagtca	600
tcgctattac	catggtgatg	cggttttggc	agtacatcaa	tgggcgtgga	tagcggtttg	660
actcacgggg	atttccaagt	ctccacccca	ttgacgtcaa	tgggagtttg	ttttggcacc	720
aaaatcaacg	ggactttcca	aaatgtcgta	acaactccgc	cccatgacg	caaatgggcg	780
gtaggcgtgt	acggtgggag	gtctatataa	gcagagctct	ctggctaact	agagaaccca	840
ctgcttactg	gcttatcgaa	attaatacga	ctcactatag	ggagacccaa	gctggctagt	900
taagcttcca	tgccggggtt	ttacgagatt	gtgattaagg	tcccagcga	ccttgacggg	960
catctgcccc	gcattttctga	cagctttgtg	aactgggttg	ccgagaagga	atgggagttg	1020
ccgccagatt	ctgacatgga	tctgaatctg	attgagcagg	cacccctgac	cgtggccgag	1080
aagctgcagc	gcgactttct	gacggaatgg	cgccgtgtga	gtaaggcccc	ggaggccctt	1140
ttctttgtgc	aatttgagaa	gggagagagc	tacttccaca	tgacgtgct	cgtggaaacc	1200
accgggggtga	aatccatggt	tttgggacgt	ttcctgagtc	agattcgcg	aaaactgatt	1260
cagagaattt	accgcgggat	cgagccgact	ttgccaaact	ggttcgcggt	cacaaagacc	1320
agaaatggcg	ccggaggcgg	gaacaaggtg	gtggatgagt	gctacatccc	caattacttg	1380
ctccccaaaa	cccagcctga	gctccagtgg	gcgtggacta	atatggaaca	gtatttaagc	1440
gcctgtttga	atctcacgga	gcgtaaacgg	ttggtggcgc	agcatctgac	gcacgtgtcg	1500
cagacgcagg	agcagaacaa	agagaatcag	aatcccaatt	ctgatgcgcc	ggtgatcaga	1560
tcaaaaactt	cagccaggta	catggagctg	gtcgggtggc	tcgtggacaa	ggggattacc	1620
tcggagaagc	agtggatcca	ggaggaccag	gcctcataca	tctccttcaa	tgccgcctcc	1680
aactcgcggt	cccaaataca	ggctgccttg	gacaatgcgg	gaaagattat	gagcctgact	1740
aaaaccgccc	ccgactacct	ggtgggccag	cagcccgtgg	aggacatttc	cagcaatcgg	1800
atttataaaa	ttttggaact	aaacgggtac	gatccccaat	atgcggcttc	cgtctttctg	1860
ggatgggcca	cgaaaaagtt	cggcaagagg	aacaccatct	ggctgttttg	gcctgcaact	1920
accgggaaga	ccaacatcgc	ggaggccata	gcccacactg	tgcccttcta	cgggtgcgta	1980
aactggacca	atgagaactt	tcccttcaac	gactgtgtcg	acaagatggt	gatctgggtg	2040
gaggagggga	agatgaccgc	caaggtcgtg	gagtcggcca	aagccattct	cggaggaagc	2100
aaggtgcgcg	tggaccagaa	atgcaagtcc	tcggcccaga	tagacccgac	tcccgatgac	2160
gtcacctcca	acaccaacat	gtgcgccgtg	attgacggga	actcaacgac	cttcgaacac	2220
cagcagccgt	tgcaagaccg	gatgttcaaa	tttgaactca	cccgcgctct	ggatcatgac	2280
tttgggaagg	tcaccaagca	ggaagtcaaa	gactttttcc	ggtgggcaaa	ggatcacgtg	2340
gttgaggtgg	agcatgaatt	ctacgtcaaa	aaggggtggag	ccgatatcca	gcacagtggc	2400
ggccgcatga	cctctcgccg	ctccgtgaag	tcgggtccgc	gggaggttcc	gcgcgatgag	2460
tacgaggatc	tgtactacac	cccgtcttca	ggtatggcga	gtcccgatag	tccgcctgac	2520
acctcccgcc	gtggcgccct	acagacacgc	tcgcgccaga	ggggcgaggt	ccgtttcgtc	2580
cagtacgacg	agtcggatta	tgccctctac	gggggctcgt	cttcgaaga	cgacgaacac	2640
ccggaggtcc	cccggacgcg	gcgtcccgtt	tcgggggcgg	ttttgtccgg	cccggggcct	2700
gcgcggggcg	ctccgccacc	cgctgggtcc	ggagggggcg	gacgcacacc	caccaccgcc	2760
ccccggggccc	ccgaaaccca	gcgggtggcg	tctaaggccc	ccgcggcccc	ggcggcggag	2820
accacccgcg	gcaggaaatc	ggcccagcca	gaatccgcgg	cactcccaga	cgcccccgcg	2880
tcgacggcgc	caaccggatc	caagacaccc	gcgcaggggc	tggccagaaa	gctgcacttt	2940
agcaccgccc	ccccaaaccc	cgacgcgcca	tggacccccc	gggtggccgg	ctttaacaag	3000
cgcgtcttct	gcgccgcggt	cgggcgcctg	gcggccatgc	atgcccggat	ggcggcggtc	3060
cagctctggg	acatgtcgcg	tccgcgcaca	gacgaagacc	tcaacgaact	ccttggcatc	3120

accaccatcc	gcgtgacggt	ctgcgagggc	aaaaacctgc	ttcagcgcg	caacgagttg	3180
gtgaatccag	acgtggtgca	ggacgtcgac	gcggccacgg	cgactcgagg	gcgttctgcg	3240
gcgtcgcgcc	ccaccgagcg	acctcgagcc	ccagcccgt	ccgcttctcg	ccccagacgg	3300
cccgtcgagg	gtctagaggg	cccgcggttc	gaacaaaaac	tcattctcaga	agaggatctg	3360
aatatgcata	ccggtcatca	tcaccatcac	cattgagttt	aaaccgcgtg	atcagcctcg	3420
actgtgcctt	ctagttgcca	gccatctgtt	gtttgcccct	cccccggtgc	ttccttgacc	3480
ctggaagggtg	ccactcccac	tgtcctttcc	taataaaatg	aggaaattgc	atcgcatgtt	3540
ctgagtaggt	gtcattctat	tctggggggg	gggggtgggg	aggacagcaa	gggggaggat	3600
tgggaagaca	atagcaggca	tgctggggat	gcggtgggct	ctatggcttc	tgaggcggaa	3660
agaaccagct	ggggctctag	gggggtatccc	cacgcgccct	gtagcggcg	attaagcgcg	3720
gcgggtgtgg	tggttacg	cagcgtgac	gctacacttg	ccagcgccct	agcgcccgct	3780
cctttcgctt	tcttcccttc	ctttctcgcc	acgttcgccc	gctttccccg	tcaagctcta	3840
aatcggggca	tccctttagg	gttcgatttt	agtgttttac	ggcacctcga	ccccaaaaaa	3900
cttgattagg	gtgatggttc	acgtagtggg	ccatgcacct	gatagacggg	ttttcgccct	3960
ttgacgttgg	agtccacgtt	ctttaatagt	ggactcttgt	tccaaactgg	aacaacactc	4020
aaccttatct	cggctctatt	ttttgattta	taagggattt	tggggatttc	ggcctattgg	4080
ttaaaaaatg	agctgattta	acaaaaattt	aacgcgaatt	aattctgtgg	aatgtgtgtc	4140
agttagggtg	tggaaagtcc	ccaggctccc	caggcaggca	gaagtatgca	aagcatgcat	4200
ctcaattagt	cagcaaccag	gtgtggaaag	tccccaggct	cccagcagg	cagaagtatg	4260
caaagcatgc	atctcaatta	gtcagcaacc	atagtcccgc	ccctaactcc	gcccattccc	4320
cccctaactc	cgcccagttc	cgcccattct	ccgccccatg	gctgactaat	tttttttatt	4380
tatgcagagg	ccgaggccgc	ctctgcctct	gagctattcc	agaagtagtg	aggaggcttt	4440
tttgagggcc	taggcttttg	caaaaagctc	ccgggagctt	gtatatccat	tttcggatct	4500
gatcaagaga	caggatgagg	atcgtttcgc	atgattgaac	aagatggatt	gcacgcagg	4560
tctccggccg	cttggttgga	gaggctattc	ggctatgact	gggcacaaca	gacaatcggc	4620
tgctctgatg	ccgccgtgtt	ccggctgtca	gcgcaggggc	gcccggttct	ttttgtcaag	4680
accgacctgt	ccggtgccct	gaatgaactg	caggacgagg	cagcgcggt	atcggtgctg	4740
gccacgacgg	gcgttccttg	cgcagctgtg	ctcgacgttg	tactgaagc	gggaagggac	4800
tggctgctat	tgggcgaagt	gccggggcag	gatctcctgt	catctcacct	tgctcctgcc	4860
gagaaagtat	ccatcatggc	tgatgcaatg	cggcggtgc	atacgcttga	tccggtacc	4920
tgcccattcg	accaccaagc	gaaacatcgc	atcgagcgag	cacgtactcg	gatggaagcc	4980
ggtcttgtcg	atcaggatga	tctggacgaa	gagcatcagg	ggctcgcgcc	agccgaactg	5040
ttccgaggc	tcaaggcgcg	catgcccgc	ggcgaggatc	tcgtcgtgac	ccatggcgat	5100
gcctgcttgc	cgaatatcat	ggtggaaaat	ggccgctttt	ctggattcat	cgactgtggc	5160
cggtggtg	tggcggaccg	ctatcaggac	atagcgttgg	ctaccgtga	tattgtgaa	5220
gagcttggcg	gcgaatgggc	tgaccgcttc	ctcgtgcttt	acggtatcgc	cgctcccgat	5280
tcgcagcgca	tcgcttctta	tcgcttcttt	gacgagttct	tctgagcggg	actctgggg	5340
tcgcgaaatg	accgaccaag	cgacgcccaa	cctgccatca	cgagatttcg	attccaccgc	5400
cgcttcttat	gaaagggttg	gcttcggaat	cgttttccgg	gacgcgggct	ggatgatcct	5460
ccagcgcggg	gatctcatgc	tggagttctt	cgcccacccc	aacttggtta	ttgcagctta	5520
taatggttac	aaataaagca	atagcatcac	aaatttcaca	aataaagcat	ttttttcact	5580
gcattctagt	tgtggtttgt	ccaaactcat	caatgtatct	tatcatgtct	gtataccgtc	5640
gacctctagc	tagagcttgg	cgtaatcatg	gtcatagctg	tttcctgtgt	gaaattgtta	5700
tccgctcaca	attccacaca	acatacgagc	cggaagcata	aagtgtaaag	cctgggggtg	5760
ctaataagtg	agctaactca	cattaattgc	gttgcgctca	ctgcccgtct	tccagtcggg	5820
aaacctgtcg	tgccagctgc	attaatgaat	cggccaacgc	gcggggagag	gcggtttg	5880
tattggggcg	tcttccgctt	cctcgctcac	tgactcgctg	cgctcggtcg	ttcggtg	5940
gcgagcggta	tcagctcact	caaaggcgg	aatacgggtta	tccacagaat	caggggataa	6000

cgcaggaaaag	aacatgtgag	caaaaaggcca	gcaaaaaggcc	aggaaccgta	aaaaggccgc	6060
gttgctggcg	tttttccata	ggctccgccc	ccctgacgag	catcacaaaa	atcgacgctc	6120
aagtcagagg	tggcgaaacc	cgacaggact	ataaagatac	caggcgtttc	cccctggaag	6180
ctccctcgtg	cgctctcctg	ttccgaccct	gccgcttacc	ggatacctgt	ccgcctttct	6240
cccttcggga	agcgtggcgc	tttctcaatg	ctcacgctgt	aggtatctca	gttcggtgta	6300
ggtcgttcgc	tccaagctgg	gctgtgtgca	cgaacccccc	gttcagcccg	accgctgcgc	6360
cttatccggt	aactatcgtc	ttgagtccaa	cccggtaaga	cacgacttat	cgccactggc	6420
agcagccact	ggtaacagga	ttagcagagc	gaggtatgta	ggcggtgcta	cagagttcct	6480
gaagtgggtg	cctaactacg	gctacactag	aaggacagta	tttggtatct	gcgctctgct	6540
gaagccagtt	accttcggaa	aaagagttgg	tagctcttga	tccggcaaac	aaaccaccgc	6600
tggtagcggg	ggtttttttg	tttgcaagca	gcagattacg	cgcagaaaaa	aaggatctca	6660
agaagatcct	ttgatctttt	ctacggggtc	tgacgctcag	tggaacgaaa	actcacgtta	6720
agggattttg	gtcatgagat	tatcaaaaag	gatcttcacc	tagatccttt	taaattaaaa	6780
atgaagtttt	aatcaatct	aaagtatata	tgagtaaact	tggctcgaca	gttaccaatg	6840
cttaatcagt	gaggcaccta	tctcagcgat	ctgtctattt	cgttcatcca	tagttgcctg	6900
actccccgtc	gtgtagataa	ctacgatacg	ggagggctta	ccatctggcc	ccagtgcctgc	6960
aatgataccg	cgagaccac	gctcaccggc	tccagattta	tcagcaataa	accagccagc	7020
cggaaagggc	gagcgcajaa	gtggtcctgc	aactttatcc	gcctccatcc	agtctattaa	7080
ttgttgccgg	gaagctagag	taagtagttc	gccagttaat	agtttgcgca	acgttggtgc	7140
cattgctaca	ggcatcgtag	gttcacgctc	gtcgtttggg	atggcttcac	tcagctccgg	7200
ttcccaacga	tcaaggcgag	ttacatgata	ccccatggtg	tgcaaaaaag	cggttagctc	7260
cttcgggtcct	ccgatcggtg	tcagaagtaa	gttgggccgca	gtgttatcac	tcatggttat	7320
ggcagcactg	cataattctc	ttactgtcat	gccatccgta	agatgctttt	ctgtgactgg	7380
tgagtactca	accaagtcat	tctgagaata	gtgtatgcgg	cgaccgagtt	gctcttgccc	7440
ggcgtcaata	cgggataata	ccgcgccaca	tagcagaact	ttaaagtgcc	tcattcattgg	7500
aaaacgttct	tcggggcgaa	aactctcaag	gatcttaccg	ctggtgagat	ccagttcgat	7560
gtaacccact	cgtgcaccca	actgatcttc	agcatctttt	actttcacca	gcgtttctgg	7620
gtgagcaaaa	acaggaaggc	aaaatgccgc	aaaaaaggga	ataagggcga	cacggaaatg	7680
ttgaatactc	atactcttcc	tttttcaata	ttattgaagc	atttatcagg	gttattgtct	7740
catgagcgga	tacatatgtg	aatgtattta	gaaaaataaa	caaatagggg	ttccgcgcac	7800
atttccccga	aaagtgccac	ctgacgctc				7828

(SEQ ID NO:22)

**Table V. RepVP22-R<sub>490</sub> Sequence**

gacggatcgg	gagatctccc	gatcccctat	ggtcgactct	cagtacaatc	tgctctgatg	60
ccgcatagtt	aagccagtat	ctgctccctg	cttggtgtgt	ggaggctcgt	gagtagtgcg	120
cgagcaaaat	ttaagctaca	acaaggcaag	gcttgaccga	caattgcatg	aagaatctgc	180
ttagggttag	gcgttttgcg	ctgcttcgcg	atgtacgggc	cagatatacg	cgttgacatt	240
gattattgac	tagttattaa	tagtaatcaa	ttacgggggc	attagttcat	agccatata	300
tggagttccg	cgttacataa	cttacggtaa	atggcccggc	tggctgaccg	cccaacgacc	360
cccgccatt	gacgtcaata	atgacgtatg	ttcccatagt	aacgccataa	gggactttcc	420
attgacgtca	atgggtggac	tatttacggg	aaactgccc	cttggcagta	catcaagtgt	480
atcatatgcc	aagtacgcc	cctattgacg	tcaatgacgg	taaatggccc	gcctggcatt	540
atgccagta	catgacctta	tgggactttc	ctacttggca	gtacatctac	gtattagtca	600
tcgctattac	catggtgatg	cggttttggc	agtacatcaa	tgggcgtgga	tagcggtttg	660
actcacgggg	atttccaagt	ctccacccca	ttgacgtcaa	tgggagtttg	ttttggcacc	720



aaaatcaacg	ggactttcca	aaatgtcgta	acaactccgc	cccattgacg	caaattggcg	780
gtaggcgtgt	acggtgggag	gtctatataa	gcagagctct	ctggctaact	agagaacca	840
ctgcttactg	gcttatcgaa	attaatacga	ctcactatag	ggagacccaa	gctggctagt	900
taagcttggt	accgagctcg	gatccactag	tccagtgtgg	tggaattctg	cagatatcat	960
gccgggggtt	tacgagattg	tgattaaggt	ccccagcgac	cttgacgggc	atctgcccgg	1020
catttctgac	agctttgtga	actgggtggc	cgagaaggaa	tgggagtgtc	cgccagattc	1080
tgacatggat	ctgaatctga	ttgagcaggc	acccctgacc	gtggccgaga	agctgcagcg	1140
cgactttctg	acggaatggc	gccgtgtgag	taaggccccg	gaggcccttt	tctttgtgca	1200
at ttgagaag	ggagagagct	acttccacat	gcacgtgctc	gtggaaacca	ccggggtgaa	1260
atccatgggt	ttgggacgtt	tcctgagtca	gattcgcgaa	aaactgattc	agagaattta	1320
ccgcgggac	gagccgactt	tgccaaactg	gttcgcggtc	acaaagacca	gaaatggcgc	1380
cggaggcggg	aacaaggtgg	tggtgagtg	ctacatcccc	aattacttgc	tccccaaaac	1440
ccagcctgag	ctccagtggg	cgtggactaa	tatggaacag	tatttaagcg	cctgtttgaa	1500
tctcacggag	cgtaaacggt	tgggtggcgca	gcatctgacg	cacgtgtcgc	agacgcagga	1560
gcagaacaaa	gagaatcaga	atcccaattc	tgatgcgccg	gtgatcagat	caaaaacttc	1620
agccaggtag	atggagctgg	tcgggtggct	cgtggacaag	gggattacct	cggagaagca	1680
gtggatccag	gaggaccagg	cctcatacat	ctccttcaat	gcggcctcca	actcgcggtc	1740
ccaaatcaag	gctgccttgg	acaatgcggg	aaagattatg	agcctgacta	aaaccgcccc	1800
cgactacctg	gtgggccagc	agcccgtgga	ggacatttcc	agcaatcgga	tttataaaat	1860
tttggaacta	aacgggtacg	atccccaata	tgcggcttcc	gtctttctgg	gatgggccac	1920
gaaaaagttc	ggcaagagga	acaccatctg	gctgtttggg	cctgcaacta	ccgggaagac	1980
caacatcgcg	gaggccatag	cccacactgt	gcccttctac	gggtgcgtaa	actggaccaa	2040
tgagaacttt	cccttcaacg	actgtgtcga	caagatggtg	atctggtggg	aggaggggaa	2100
gatgaccgcc	aaggtcgtgg	agtcggccaa	agccattctc	ggaggaagca	aggtgcgcgt	2160
ggaccagaaa	tgcaagtcc	cggccagat	agaccgact	cccgtgatcg	tcacctcaa	2220
caccaacatg	tgcgcgtga	ttgacgggaa	ctcaacgacc	ttcgaacacc	agcagccgtt	2280
gcaagaccgg	atgttcaaat	ttgaactcac	ccgcgctctg	gatcatgact	ttgggaaggt	2340
caccaagcag	gaagtcaaa	actttttccg	gtgggcaaag	gatcacgtgg	ttgaggtgga	2400
gcatgaattc	tacgtcaaaa	aggggtggcg	ccgcatgacc	tctcgcgct	ccgtgaagtc	2460
gggtccgcgg	gaggttccgc	gcatgagta	cgaggatctg	tactacaccc	cgtcttcagg	2520
tatggcgagt	cccgatagtc	cgcctgacac	ctcccgccgt	ggcgccctac	agacacgctc	2580
gcgccagagg	ggcgagggtc	gtttcgtcca	gtacgacgag	tcggattatg	ccctctacgg	2640
gggctcgtct	tccgaagacg	acgaacaccc	ggaggtcccc	cggacgcggc	gtcccgtttc	2700
cggggcggtt	ttgtccggcc	cggggcctgc	gcgggcgcct	ccgccacccg	ctgggtccgg	2760
aggggccgga	cgcacaccca	ccaccgcccc	ccgggcccc	cgaaccacgc	gggtggcgtc	2820
taaggccccc	gcggccccgg	cggcgagac	caccgcggc	aggaaatcgg	cccagccaga	2880
atccgcgcga	ctccagacg	ccccgcgtc	gacggcgcca	acccgatcca	agacacccgc	2940
gcaggggctg	gccagaaagc	tgcactttag	caccgcccc	ccaaaccccg	acgcgccatg	3000
gaccccccg	gtggccggct	ttaacaagcg	cgtcttctgc	gccgcggctg	ggcgccctggc	3060
ggccatgcat	gcccggatgg	cggcggtcca	gctctgggac	atgtcgcgtc	cgcgcacaga	3120
cgaagacctc	aacgaactcc	ttggcatcac	caccatccgc	gtgacgggtc	gcgagggcaa	3180
aaacctgctt	cagcgcgcca	acgagttggt	gaatccagac	gtggtgcagg	acgtcgacgc	3240
ggccaacggc	actcgagggc	gttctgcggc	gtcgcgcccc	accgagcgac	ctcgagcccc	3300
agcccgcctc	gcttctcgcc	ccagacggcc	cgtcgagggt	ctagaggggc	cgcggttcga	3360
acaaaaactc	atctcagaag	aggatctgaa	tatgcatacc	ggtcatcatc	accatcacca	3420
ttgagtttaa	acccgctgat	cagcctcgac	tgtgccttct	agttgccagc	catctgttgt	3480
ttgcccctcc	cccgtgcctt	ccttgaccct	ggaaggtgcc	actcccactg	tcctttccta	3540
ataaaatgag	gaaattgcat	cgcattgtct	gagtaggtgt	cattctattc	tgggggggtg	3600

ggtggggcag	gacagcaagg	gggaggattg	ggaagacaat	agcaggcatg	ctggggatgc	3660
ggtgggctct	atggcttctg	aggcggaag	aaccagctgg	ggctctaggg	ggtatcccca	3720
cgcgccctgt	agcggcgc	taagcgcg	gggtgtgtg	gttacgcgca	gcgtgaccgc	3780
tacacttgcc	agcgccctag	cgcccgctcc	tttcgctttc	ttcccttcct	ttctcgccac	3840
gttcgcgggc	tttccccgtc	aagctctaaa	tcggggcatc	cctttagggg	tccgatttag	3900
tgctttacgg	cacctcgacc	caaaaaaact	tgattagggg	gatggttcac	gtagtgggcc	3960
atcgccctga	tagacggttt	ttcgcccttt	gacgttgagg	tccacgttct	ttaatagtgg	4020
actcttggtc	caaactggaa	caacactcaa	ccctatctcg	gtctattctt	ttgattttata	4080
agggattttg	gggatttcgg	cctattgggt	aaaaaatgag	ctgatttaac	aaaaatttaa	4140
cggaatttaa	ttctgtggaa	tgtgtgtcag	ttaggggtgtg	gaaagtcccc	agggtcccca	4200
ggcaggcaga	agtatgcaaa	gcatgcatct	caattagtca	gcaaccaggt	gtggaaagtc	4260
cccaggctcc	ccagcaggca	gaagtatgca	aagcatgcat	ctcaattagt	cagcaaccat	4320
agtcccgccc	ctaactccgc	ccatcccgcc	cctaactccg	cccagttccg	cccattctcc	4380
gccccatggc	tgactaat	tttttattta	tgagaggcc	gaggccgcct	ctgcctctga	4440
gctattccag	aagtagtgag	gaggcttttt	tggaggccta	ggcttttgca	aaaagctccc	4500
gggagcttgt	atatccattt	tcggatctga	tcaagagaca	ggatgaggat	cgtttcgcat	4560
gattgaacaa	gatggattgc	acgcagggtc	tccggccgct	tgggtggaga	ggctattcgg	4620
ctatgactgg	gcacaacaga	caatcggtg	ctctgatgcc	gccgtgttcc	ggctgtcagc	4680
gcaggggcgc	ccggttcttt	ttgtcaagac	cgacctgtcc	ggtgccctga	atgaactgca	4740
ggacgaggca	gcgcggctat	cgtggctggc	cacgacgggc	gttccttgcg	cagctgtgct	4800
cgacgttgct	actgaagcgg	gaagggaactg	gctgctattg	ggcgaagtgc	cggggcagga	4860
tctcctgtca	tctcaccttg	ctcctgccga	gaaagtatcc	atcatggctg	atgcaatgcg	4920
gcggctgcat	acgcttgatc	cggctacctg	cccattcgac	caccaagcga	aacatcgcat	4980
cgagcgagca	cgtactcgga	tggaaagccg	tcttgctgat	caggatgatc	tggacgaaga	5040
gcatcagggg	ctcgcgccag	ccgaactgtt	cgccaggctc	aaggcgcgca	tgcccgacgg	5100
cgaggatctc	gtcgtgaccc	atggcgatgc	ctgcttgccg	aatatcatgg	tggaaaatgg	5160
ccgcttttct	ggattcatcg	actgtggccg	gctgggtgtg	gcggaccgct	atcaggacat	5220
agcgttggct	acccgtgata	ttgctgaaga	gcttggcggc	gaatgggctg	accgcttctt	5280
cgtgctttac	ggtatcgccg	ctcccgattc	gcagcgcatc	gccttctatc	gccttcttga	5340
cgagttcttc	tgagcgggac	tctgggggtc	gcgaaatgac	cgaccaagcg	acgcccaccc	5400
tgccatcacg	agatttcgat	tccaccgccg	ccttctatga	aaggttgggc	ttcggaatcg	5460
ttttccggga	cgccggctgg	atgatcctcc	agcgcgggga	tctcatgctg	gagttcttcg	5520
cccaccccaa	cttggttatt	gcagcttata	atgggttaca	ataaagcaat	agcatcacia	5580
atttcacaaa	taaagcattt	ttttcactgc	attctagttg	tggtttgtcc	aaactcatca	5640
atgtatctta	tcatgtctgt	ataccgtcga	cctctagcta	gagcttggcg	taatcatggt	5700
catagctggt	tctgtgtga	aattgttatc	cgctcacaat	tccacacaac	atacgagccg	5760
gaagcataaa	gtgtaaagcc	tgggggtgct	aatgagttag	ctaactcaca	ttaattgcgt	5820
tgcgctcact	gcccgccttc	cagtcgggaa	acctgtcgtg	ccagctgcat	taatgaatcg	5880
gccaacgcgc	ggggagaggc	ggtttgctga	ttgggcgctc	ttccgcttcc	tcgctcactg	5940
actcgctgcg	ctcggtcggt	cggctgcggc	gagcggtatc	agctcactca	aaggcggtaa	6000
tacggttatc	cacagaatca	ggggataacg	caggaaagaa	catgtgagca	aaaggccagc	6060
aaaaggccag	gaaccgtaaa	aaggccgcgt	tgctggcggt	tttccatagg	ctccgcccc	6120
ctgacgagca	tcacaaaaat	cgacgtctaa	gtcagagggtg	gcgaaacccg	acaggactat	6180
aaagatacca	ggcggttccc	cctggaagct	ccctcggtcg	ctctcctgtt	ccgaccctgc	6240
cgcttaccgg	atacctgtcc	gcctttctcc	cttcgggaag	cgtggcgctt	tctcaatgct	6300
cacgctgtag	gtatctcagt	tcggtgtagg	tcgttcgctc	caagctgggc	tgtgtgcacg	6360
aaccccccg	tcagcccgac	cgctgcgcct	tatccggtaa	ctatcgctct	gagtccaacc	6420
cggtaagaca	cgacttatcg	ccactggcag	cagccactgg	taacaggatt	agcagagcga	6480

ggtatgtagg	cgggtgctaca	gagttcttga	agtgggtggcc	taactacggc	tacactagaa	6540
ggacagtatt	tggtatctgc	gctctgctga	agccagttac	cttcggaaaa	agagttggta	6600
gctcttgatc	cggcaaaaca	accaccgctg	gtagcgggtg	tttttttggt	tgcaagcagc	6660
agattacgcg	cagaaaaaaa	ggatctcaag	aagatccttt	gatcttttct	acggggtctg	6720
acgctcagtg	gaacgaaaac	tcacgttaag	ggattttggt	catgagatta	tcaaaaagga	6780
tcttcaccta	gatcctttta	aattaaaaat	gaagttttaa	atcaatctaa	agtatatatg	6840
agtaaacttg	gtctgacagt	taccaatgct	taatcagtg	ggcacctatc	tcagcgatct	6900
gtctatttgc	ttcatccata	gttgccctgac	tccccgtcgt	gtagataact	acgatacggg	6960
agggcttacc	atctggcccc	agtgtctgcaa	tgataccgcg	agacccacgc	tcaccggctc	7020
cagattttatc	agcaataaac	cagccagccg	gaagggccga	gcgcagaagt	ggtcctgcaa	7080
ctttatccgc	ctccatccag	tctattaatt	gttgccggga	agctagagta	agtagttcgc	7140
cagttaatag	tttgcgcaac	gttggttgcca	ttgctacagg	catcgtgggtg	tcacgctcgt	7200
cgtttggtat	ggcttcattc	agctccgggt	cccaacgatc	aaggcgagtt	acatgatccc	7260
ccatgttggtg	caaaaaagcg	gttagctcct	tcggctcctc	gatcgttggtc	agaagtaagt	7320
tgccgcagtg	gttatcactc	atgggttatgg	cagcactgca	taattctctt	actgtcatgc	7380
catccgtaag	atgcttttct	gtgactgggtg	agtactcaac	caagtcattc	tgagaatagt	7440
gtatgcggcg	accgagttgc	tcttgcccg	cgtcaatacg	ggataatacc	gcgccacata	7500
gcagaacttt	aaaagtgtc	atcattggaa	aacgttcttc	ggggcgaaaa	ctctcaagga	7560
tcttacgct	gttgagatcc	agttcgatgt	aaccactcg	tgacccaac	tgatcttcag	7620
catcttttac	tttcaccagc	gtttctgggt	gagcaaaaac	aggaaggcaa	aatgccgcaa	7680
aaaagggaat	aagggcgaca	cggaaatgtt	gaatactcat	actcttcctt	tttcaatatt	7740
attgaagcat	ttatcagggg	tattgtctca	tgagcggata	catatttgaa	tgtatttaga	7800
aaaataaaca	aataggggtt	ccgcgcacat	ttccccgaaa	agtgccacct	gacgtc	7856

(SEQ ID NO:23)

**Table VI. VP22#2-RepVP5-EcRV-S-Cys Sequence.**

gacggatcgg	gagatctccc	gatcccctat	ggtcgactct	cagtacaatc	tgctctgatg	60
ccgcatagtt	aagccagtat	ctgctccctg	cttggtgtgt	ggaggtcgct	gagtagtgcg	120
cgagcaaaat	ttaagctaca	acaaggcaag	gcttgaccga	caattgcatg	aagaatctgc	180
ttagggttag	gcgttttgcg	ctgcttcgcg	atgtacgggc	cagatatacg	cgttgacatt	240
gattattgac	tagttattaa	tagtaatcaa	ttacgggggtc	attagttcat	agcccatata	300
tgaggttccg	cgttacataa	cttacggtaa	atggcccggc	tggtgaccg	ccaacgacc	360
cccgccatt	gacgtcaata	atgacgtatg	ttcccatagt	aacgcccaata	gggactttcc	420
attgacgtca	atgggtggac	tatttacggt	aaactgcca	cttggcagta	catcaagtgt	480
atcatatgcc	aagtacgccc	cctattgacg	tcaatgacgg	taaatggccc	gcctggcatt	540
atgccagta	catgacctta	tgggactttc	ctacttgcca	gtacatctac	gtattagtca	600
tcgctattac	catggtgatg	cggttttggc	agtacatcaa	tgggcgtgga	tagcggtttg	660
actcacgggg	atttccaagt	ctccacccca	ttgacgtcaa	tgggagtttg	ttttggcacc	720
aaaatcaacg	ggactttcca	aaatgtcgt	acaactccgc	cccattgacg	caaattggcg	780
gtaggcgtgt	acggtgggag	gtctatataa	gcagagctct	ctggctaact	agagaacca	840
ctgcttactg	gcttatcgaa	attaatacga	ctcactatag	ggagacccaa	gctggctagt	900
taagcttggt	accgagctcg	gatccactag	tccagtgtgg	tgggaattctg	cagatatcat	960
gccggggttt	tacgagattg	tgattaaggt	ccccagcgac	cttgacgggc	atctgcccg	1020
catttctgac	agctttgtga	actgggtggc	cgagaaggaa	tgggagttgc	cgccagattc	1080
tgacatggat	ctgaatctga	ttgagcaggc	accctgacc	gtggccgaga	agctgcagcg	1140
cgactttctg	acggaatggc	gccgtgtgag	taaggccccg	gaggcccttt	tctttgtgca	1200

at ttgagaag	ggagagagct	acttccacat	gcacgtgctc	gtggaaacca	ccggggtgaa	1260
atccatgggt	ttgggacgtt	tcctgagtca	gattcgcgaa	aaactgattc	agagaattta	1320
ccgcgggata	gagccgactt	tgccaaactg	gttcgcggtc	acaaagacca	gaaatggcgc	1380
cggaggcggg	aacaaggtgg	tgatgagtg	ctacatcccc	aattacttgc	tccccaaaac	1440
ccagcctgag	ctccagtggg	cgtggactaa	tatggaacag	tatttaagcg	cctgtttgaa	1500
tctcacggag	cgtaaacggt	tgggtggcga	gcacgtgacg	cacgtgtcgc	agacgcagga	1560
gcagaacaaa	gagaatcaga	atcccaattc	tgatgcgccg	gtgatcagat	caaaaacttc	1620
agccagggtac	atggagctgg	tcgggtggct	cgtggacaag	gggattacct	cggagaagca	1680
gtggatccag	gaggaccagg	cctcatatat	ctccttcaat	gcggcctcca	actcgcggtc	1740
ccaaatcaag	gctgccttgg	acaatgcggg	aaagattatg	agcctgacta	aaaccgcccc	1800
cgactacctg	gtgggccagc	agcccggtga	ggacatttcc	agcaatcgga	tttataaaat	1860
tttggaacta	aacgggtacg	atccccaata	tgccggttcc	gtctttcttg	gatgggccac	1920
gaaaaagtgc	ggcaagagga	acaccatctg	gctgtttggg	cctgcaacta	ccgggaagac	1980
caacatcgcg	gaggccatag	cccacactgt	gcccttctac	gggtgcgtaa	actggacca	2040
tgagaacttt	cccttcaacg	actgtgtcga	caagatggtg	atctggtggg	aggaggggaa	2100
gatgaccgcc	aaggtcgtgg	agtcggccaa	agccattctc	ggaggaagca	aggtgcgcgt	2160
ggaccagaaa	tgcaagtcc	cggcccgat	agaccgact	cccgatgacg	tcacctccaa	2220
caccaacatg	tgcccggtga	ttgacgggaa	ctcaacgacc	ttcgaacacc	agcagccgtt	2280
gcaagaccgg	atgttcaaat	ttgaactcac	ccgccgtctg	gatcatgact	ttgggaaggt	2340
caccaagcag	gaagtcaaa	actttttccg	gtgggcaaa	gatcacgtgg	ttgaggtgga	2400
gcatgaattc	tacgtcaaaa	aggttgagc	caagaaaaga	cccgcctcca	gtgacgcaga	2460
tataagttag	cccaaacggg	tgcccgagtc	agttgcgcag	ccatcgacgt	cagacgcgga	2520
agcttcgatc	aactacgcag	acaggtacca	aaacaaatgt	tctcgtcacg	tgggcatgaa	2580
tctgatgctg	tttccctgca	gacaatgcga	gagaatgaat	cagaattcaa	atatctgctt	2640
cactcacgga	cagaaagact	gttttagagt	ctttcccggt	tcagaatctc	aaccggtttc	2700
tgctgtcaaa	aaggcgtatc	agaaactgtg	ctacattcat	catatcatgg	gaaaggtgcc	2760
agacgcttgc	actgcctgcg	atctggtcaa	tgtggatttg	gatgactgca	tctttgaaca	2820
atgcggccgc	atgacctctc	gccgtccgt	gaagtccggg	ccgcgggagg	ttccgcgcga	2880
tgagtacgag	gatctgtact	acacccgctc	ttcaggatag	gcgagtcctg	atagtccgcc	2940
tgacacctcc	cgcctggcgg	ccctacagac	acgctcgcgc	cagaggggag	aggtccgttt	3000
cgctccagtac	gacgagtcgg	attatgccct	ctacgggggc	tcgtcttccg	aagacgcagc	3060
acacccgagg	gtcccccgga	cgcggcgtcc	cgtttccggg	gcggttttgt	ccggcccggg	3120
gcctgcgcgg	gcgcctccgc	caccgctgg	gtccggaggg	gccggacgca	caccaccac	3180
cgcctcccg	gcccccgaa	cccagcgggt	ggcgtctaag	gccccgcgg	ccccggcggc	3240
ggagaccacc	cgcggcagga	aatcggccca	gccagaatcc	gccgcactcc	cagacgcccc	3300
cgcgtcgacg	gcgccaaccc	gatccaagac	acccgcgcag	gggtggcca	gaaagctgca	3360
cttttagcacc	gcccccccaa	accccgacgc	gccatggacc	ccccgggtgg	ccggctttaa	3420
caagcgcgct	ttctgcgccg	cggtcggggc	cctggcggcc	atgcatgccc	ggatggcggc	3480
ggtccagctc	tgggacatgt	cgcgtcccg	cacagacgaa	gacctcaacg	aactccttgg	3540
catcaccacc	atccgcgtga	cggctctcga	gggcaaaaac	ctgcttcagc	gcgccaacga	3600
gttggtgaat	ccagacgtgg	tgaggacgt	cgacgcggcc	acggcgactc	gagggcgttc	3660
tgccgctcgc	cgcctcccg	agcgacctcg	agccccagcc	cgctccgctt	ctcgccccag	3720
acggcccgct	gagggtctag	agggcccgcg	gttcgaacaa	aaactcatct	cagaagagga	3780
tctgaatatg	cataccggtc	atcatcacca	tcaccattga	gtttaaaccc	gctgatcagc	3840
ctcgactgtg	ccttctagtt	gccagccatc	tggtgtttgc	ccctcccccg	tgcttccctt	3900
gaccttgaa	ggtgccactc	ccactgtcct	ttcctaataa	aatgaggaaa	ttgcatcgca	3960
ttgtctgagt	aggtgtcatt	ctattctggg	gggtgggggtg	gggcaggaca	gcaaggggga	4020
ggattgggaa	gacaatagca	ggcatgctgg	ggatgcgggtg	ggctctatgg	cttctgaggc	4080

ggaaagaacc	agctggggct	ctagggggta	tccccacgcg	ccctgtagcg	gcgcattaag	4140
cgcggcgggt	gtggtggtta	cgcgcagcgt	gaccgctaca	cttgccagcg	ccctagcgcc	4200
cgctcctttc	gctttcttcc	cttcctttct	cgccacgttc	gccggctttc	cccgtcaagc	4260
tctaaatcgg	ggcatccctt	tagggttccg	atttagtgct	ttacggcacc	tcgaccccaa	4320
aaaacttgat	tagggtgatg	gttcacgtag	tgggccatcg	ccctgataga	cggtttttcg	4380
ccctttgacg	ttggagtcca	cgttctttta	tagtggactc	ttgttccaaa	ctggaacaac	4440
actcaaccct	atctcggtct	attcttttga	tttataaggg	attttgggga	tttcggccta	4500
ttggttaaaa	aatgagctga	tttaacaaaa	atttaacgcg	aattaattct	gtggaatgtg	4560
tgtcagttag	ggtgtggaaa	gtccccaggc	tccccaggca	ggcagaagta	tgcaaagcat	4620
gcatctcaat	tagtcagcaa	ccaggtgtgg	aaagtcccca	ggctccccag	caggcagaag	4680
tatgcaaagc	atgcatctca	attagtcagc	aaccatagtc	ccgcccctaa	ctccgcccac	4740
cccgcacctc	actccgcccc	gttcgcgcca	ttctccgccc	catggctgac	taattttttt	4800
tattttatgca	gaggccgagg	ccgcctctgc	ctctgagcta	ttccagaagt	agtgaggagg	4860
cttttttggg	ggcctagggt	tttgcaaaaa	gctcccggga	gcttgatat	ccatttttcg	4920
atctgatcaa	gagacaggat	gaggatcggt	tcgcatgatt	gaacaagatg	gattgcacgc	4980
aggttctccg	gccgcttggg	tggagaggct	attcggctat	gactgggcac	aacagacaat	5040
cggtctctct	gatgccgccc	tggtccggct	gtcagcgccg	gggcgcccgg	ttctttttgt	5100
caagaccgac	ctgtccggtg	ccctgaatga	actgcaggac	gaggcagcgc	ggctatcgtg	5160
gctggccacg	acgggcgttc	cttgccgagc	tgtgctcgac	gttgctactg	aagcgggaag	5220
ggactggctg	ctattgggcg	aagtgccggg	gcaggatctc	ctgtcatctc	accttgctcc	5280
tgccgagaaa	gtatccatca	tggctgatgc	aatgcggcgg	ctgcatacgc	ttgatccggc	5340
tacctgcccc	ttcgaccacc	aagcgaaaca	tcgcatcgag	cgagcacgta	ctcggatgga	5400
agccgggtctt	gtcgatcagg	atgatctgga	cgaagagcat	caggggctcg	cgccagccga	5460
actgttcgcc	aggctcaagg	cgcgcatgcc	cgacggcgag	gatctcgctg	tgacccatgg	5520
cgatgcctgc	ttgccgaata	tcatggtgga	aaatggccgc	ttttctggat	tcacgcactg	5580
tggccggctg	ggtgtggcgg	accgctatca	ggacatagcg	ttggctaccc	gtgatattgc	5640
tgaagagctt	ggcggcgaat	gggctgaccg	cttcctcgct	ctttacggta	tcgccgctcc	5700
cgattcgcag	cgcacgcctt	tctatcgctt	tcttgacgag	ttcttctgag	cgggactctg	5760
gggttcgcga	aatgaccgac	caagcgacgc	ccaacctgcc	atcacgagat	ttcgattcca	5820
ccgccgcctt	ctatgaaagg	ttgggcttcg	gaatcgtttt	ccgggacgcc	ggctggatga	5880
tctccagcg	cggggatctc	atgctggagt	tcttcgcccc	ccccaaactg	tttattgcag	5940
cttataatgg	ttacaaataa	agcaatagca	tcacaaatth	cacaaataaa	gcattttttt	6000
cactgcattc	tagttgtggt	ttgtccaaac	tcacaaatth	atcttatcat	gtctgtatac	6060
cgctgacctc	tagctagagc	ttggcgtaat	catggtcata	gctgtttcct	gtgtgaaatt	6120
gttatccgct	cacaattcca	cacaacatac	gagccggaag	cataaagtgt	aaagcctggg	6180
gtgcctaata	agttagctaa	ctcacattaa	ttgcgttgcg	ctcactgcc	gctttccagt	6240
cgggaaacct	gtcgtgccag	ctgcattaat	gaatcggcc	acgcgcgggg	agaggcggtt	6300
tgcgtattgg	gcgctcttcc	gcttcctcgc	tactgactc	gctgcgctcg	gtcgttcggc	6360
tgcggcgagc	ggtatcagct	cactcaaagg	cggtaatacg	gttatccaca	gaatcagggg	6420
ataacgcagg	aaagaacatg	tgagcaaaa	gccagcaaaa	ggccaggaac	cgtaaaaagg	6480
ccgcgttgct	ggcgtttttc	cataggctcc	gccccctga	cgagcatcac	aaaaatcgac	6540
gctcaagtca	gaggtggcga	aacccgacag	gactataaag	ataccaggcg	tttccccctg	6600
gaagctccct	cgtgcgctct	cctgttccga	ccctgccgct	taccggatac	ctgtccgcct	6660
ttctcccttc	gggaagcggt	gcgctttctc	aatgctcacg	ctgtaggtat	ctcagttcgg	6720
tgtaggctgt	tcgctccaag	ctgggctgtg	tgcacgaacc	ccccgttcag	cccgaaccgt	6780
gcgccttata	cggtaactat	cgtcttgagt	ccaaccgggt	aagacacgac	ttatcgccac	6840
tggcagcagc	cactggtaac	aggattagca	gagcgaggta	tgtaggcggt	gctacagagt	6900
tcttgaagtg	gtggcctaac	tacggctaca	ctagaaggac	agtatttggt	atctgcgctc	6960

tgctgaagcc	agttaccttc	ggaaaaagag	ttggtagctc	ttgatccggc	aaacaaacca	7020
ccgctggtag	cgggtggtttt	tttgtttgca	agcagcagat	tacgcgcaga	aaaaaaggat	7080
ctcaagaaga	tcctttgata	ttttctacgg	ggctctgacgc	tcagtgggaa	gaaaactcac	7140
gttaagggat	tttggtcata	agattatcaa	aaaggatctt	cacctagatc	cttttaaatt	7200
aaaaatgaag	ttttaaatca	atctaaagta	tatatgagta	aacttggtct	gacagttacc	7260
aatgcttaat	cagtgaaggca	cctatctcag	cgatctgtct	atttcgttca	tccatagttg	7320
cctgactccc	cgtcgtgtag	ataactacga	tacgggaggg	cttaccatct	ggccccagtg	7380
ctgcaatgat	accgcgagac	ccacgctcac	cggtccaga	tttatcagca	ataaaccagc	7440
cagccggaag	ggccgagcgc	agaagtggtc	ctgcaacttt	atccgcctcc	atccagtcta	7500
ttaattgttg	ccgggaagct	agagtaagta	gttcgccagt	taatagtttg	cgcaacgttg	7560
ttgccattgc	tacaggcatc	gtggtgtcac	gctcgtcgtt	tggtatggct	tcattcagct	7620
ccggttccca	acgatcaagg	cgagttacat	gatccccc	gttgtgcaaa	aaagcggtta	7680
gctccttcgg	tcctccgatc	gttgtcagaa	gtaagtggc	cgcagtgtta	tcactcatgg	7740
ttatggcagc	actgcataat	tctcttactg	tcatgccatc	cgtaagatgc	ttttctgtga	7800
ctggtgagta	ctcaaccaag	tcattctgag	aatagtgtat	gcggcgaccg	agttgctctt	7860
gcccggcgtc	aatacgggat	aataccgcgc	cacatagcag	aactttaaaa	gtgctcatca	7920
ttggaaaacg	ttcttcgggg	cgaaaactct	caaggatctt	accgctgttg	agatccagtt	7980
cgatgtaacc	cactcgtgca	cccaactgat	cttcagcatc	ttttactttc	accagcgttt	8040
ctgggtgagc	aaaaacagga	aggcaaatg	ccgcaaaaaa	gggaataagg	gcgacacgga	8100
aatgttgaat	actcatactc	ttcctttttc	aatattattg	aagcatttat	cagggttatt	8160
gtctcatgag	cggatacata	tttgaatgta	tttagaaaaa	taaacaaata	ggggttccgc	8220
gcacatttcc	ccgaaaagtg	ccacctgacg	tc			8252

(SEQ ID NO:24)